





```

: SOFTWARE: Patentlin version 3.1
: SEQ ID NO 1
:
: LENGTH: 2443
:
: TYPE: DNA
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(2443)
:
: OTHER INFORMATION: SBFI modified region
US-60-361-534-1

```

Query Match	9.0%	Score	90.4	DB	8	Length	2443
Best Local Similarity	58.2%	Pred	NO	1.5e-11			
Matches	178	Conservative	0	Mismatches	126	Indels	2
						Gaps	1

[illegible]

```

1      RESULT      6
2      US-10-056-454A-17
3      : Sequence 17, Application US/10056454A
4      :
5      : GENERAL INFORMATION:
6      : APPLICANT: National starch and Chemical Investment Holding Corporation
7      : TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition
8      : NUMBER OF SEQUENCES: 20
9      : CORRESPONDENCE ADDRESS:
10     : ADDRESS: National starch and Chemical Investment Holding Corporation
11     : STREET: 1000 Unigema Blvd.
12     : CITY: Newcastle
13     : STATE: Delaware
14     : COUNTRY: United States of America
15     : ZIP: 19720
16     :
17     : COMPUTER READABLE FORM:
18     : MEDIUM TYPE: Floppy disk
19     : COMPUTER: IBM PC compatible
20     : OPERATING SYSTEM: PC-DOS/MS-DOS
21     : SOFTWARE: PatentIn Release #1.0, Version #1.30
22     : CURRENT APPLICATION DATA:
23     : APPLICATION NUMBER: US/10/056,454A
24     : FILING DATE: 25-Jun-2002
25     :
26     : INFORMATION FOR SEQ ID NO: 17:
27     :
28     : SEQUENCE CHARACTERISTICS:
29     :
30     : LENGTH: 2529 base pairs
31     : TYPE: nucleic acid
32     : STRANDEDNESS: single
33     : TOPOLOGY: linear
34     :
35     : SEQUENCE DESCRIPTION: SEQ ID NO: 17:
36     :
37     : US-10-056-454A-17

```

Query Match	8.8%;	Score 88;	DB 7;	Length 2529;
Best Local Similarity	73.7%;	Pred. No. 5.5e-11;		
Matches 112;	Conservative	0;	Mismatches 40;	Indels 0;
			Gaps	0;

QY 598 gctgtgtccaggtacacatgtaactaatcttttggccccaagtagcgttttggaaactca 657  
Db 1100 GCTGTGTTTTGGTATCATCTATGTCACAAAAATTTTTTGGCACCAAGAGCGTTTGGAAACGGCC 1159  
QY 658 gaggaacttaaaatccatgcatacgatagacacatgaactggtttgctgttcttaagat 717  
Db 1160 GAGCAACCTTAAAGTCTTGTGATTGTGATTAAGCCATGACGACTGAGAAATGTTGTCATGAGAC 1219  
QY 718 atgtcatagatgaatattagatccaatthaatt 749  
Db 1220 ATTGTTCACAGCCATGCATCAATAATTAATCTTT 1251

RESULT 7  
US-10-056-454A-16  
Sequence 16, Application US/10056454A  
GENERAL INFORMATION:  
APPLICANT: National Starch and Chemical Investment Holding Corporation  
TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESS: National Starch and Chemical Investment Holding Corporation  
STREET: 1000 Unigema Blvd.  
CITY: Newcastle  
STATE: Delaware  
COUNTRY: United States of America  
ZIP: 19720  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/056,454A  
FILING DATE: 25-Jun-2002  
INFORMATION FOR SEQ. ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2576 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ. ID NO: 16:  
US-10-056-454A-16

```

Query Match Similarity      8.8%; Score 88; DB 7; Length 2576;
Best Local Similarity      73.7%; Pred. No. 5.5e-11;
Matches 112; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 598 gctgttcacaggtacatgttacttaatttttgcaccaagtagccgttttggnaactcca 657
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1105 gctactttttggtttancatcgtcacaaattttttggcaccagacgccgttttggacgccc 1164

QY 658 gaggacctaataatccttgcatagcatagcatagcattggtttgtctgtttcttatgcat 717
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1165 gacgcacctttaaactcttggattgaataaacgtcagaccttgggaattgtttcycatggac 1224

QY 718 attgtcatagtgtaattagtcacattaatc 749
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1225 attgttcacagccatgcatacmaataatacttt 1256

RESULT 8
US-10-056-454A-19
; Sequence 19, Application US/10056454A
; GENERAL INFORMATION:
; APPLICANT: National Starch and Chemical Investment Holding Corporation
; TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition
; NUMBER OF SEQUENCES: 20

```













GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 07:30:52 ; Search time 5258.2 Seconds  
(without alignments)  
4114.507 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_1\_1000  
Perfect score: 1000  
Sequence: 1 agaacacctccatttaga.....aacatcccggtggtccct 1000

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 1081749327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/pdata/1/pna/PCTUS\_COMB.seq:\*  
2: /cgn2\_6/pdata/1/pna/US05\_COMB.seq:\*  
3: /cgn2\_6/pdata/1/pna/US07\_COMB.seq:\*  
4: /cgn2\_6/pdata/1/pna/US080\_COMB.seq:\*  
5: /cgn2\_6/pdata/1/pna/US081\_COMB.seq:\*  
6: /cgn2\_6/pdata/1/pna/US082\_COMB.seq:\*  
7: /cgn2\_6/pdata/1/pna/US083\_COMB.seq:\*  
8: /cgn2\_6/pdata/1/pna/US084\_COMB.seq:\*  
9: /cgn2\_6/pdata/1/pna/US085\_COMB.seq:\*  
10: /cgn2\_6/pdata/1/pna/US086\_COMB.seq:\*  
11: /cgn2\_6/pdata/1/pna/US087\_COMB.seq:\*  
12: /cgn2\_6/pdata/1/pna/US088\_COMB.seq:\*  
13: /cgn2\_6/pdata/1/pna/US089\_COMB.seq:\*  
14: /cgn2\_6/pdata/1/pna/US090\_COMB.seq:\*  
15: /cgn2\_6/pdata/1/pna/US091\_COMB.seq:\*  
16: /cgn2\_6/pdata/1/pna/US092\_COMB.seq:\*  
17: /cgn2\_6/pdata/1/pna/US093\_COMB.seq:\*  
18: /cgn2\_6/pdata/1/pna/US094\_COMB.seq:\*  
19: /cgn2\_6/pdata/1/pna/US095A\_COMB.seq:\*  
20: /cgn2\_6/pdata/1/pna/US095B\_COMB.seq:\*  
21: /cgn2\_6/pdata/1/pna/US095C\_COMB.seq:\*  
22: /cgn2\_6/pdata/1/pna/US095D\_COMB.seq:\*  
23: /cgn2\_6/pdata/1/pna/US096A\_COMB.seq:\*  
24: /cgn2\_6/pdata/1/pna/US096B\_COMB.seq:\*  
25: /cgn2\_6/pdata/1/pna/US096C\_COMB.seq:\*  
26: /cgn2\_6/pdata/1/pna/US096D\_COMB.seq:\*  
27: /cgn2\_6/pdata/1/pna/US096E\_COMB.seq:\*  
28: /cgn2\_6/pdata/1/pna/US097A\_COMB.seq:\*  
29: /cgn2\_6/pdata/1/pna/US097B\_COMB.seq:\*  
30: /cgn2\_6/pdata/1/pna/US097C\_COMB.seq:\*  
31: /cgn2\_6/pdata/1/pna/US098A\_COMB.seq:\*  
32: /cgn2\_6/pdata/1/pna/US098B\_COMB.seq:\*  
33: /cgn2\_6/pdata/1/pna/US098C\_COMB.seq:\*  
34: /cgn2\_6/pdata/1/pna/US099A\_COMB.seq:\*  
35: /cgn2\_6/pdata/1/pna/US099B\_COMB.seq:\*  
36: /cgn2\_6/pdata/1/pna/US099C\_COMB.seq:\*  
37: /cgn2\_6/pdata/1/pna/US100\_COMB.seq:\*  
38: /cgn2\_6/pdata/1/pna/US101\_COMB.seq:\*  
39: /cgn2\_6/pdata/1/pna/US6000\_COMB.seq:\*  
40: /cgn2\_6/pdata/1/pna/US6001\_COMB.seq:\*  
41: /cgn2\_6/pdata/1/pna/US6002\_COMB.seq:\*  
42: /cgn2\_6/pdata/1/pna/US6003\_COMB.seq:\*  
43: /cgn2\_6/pdata/1/pna/US6004\_COMB.seq:\*

44: /cgn2\_6/pdata/1/pna/US6005\_COMB.seq:\*  
45: /cgn2\_6/pdata/1/pna/US6006\_COMB.seq:\*  
46: /cgn2\_6/pdata/1/pna/US6007\_COMB.seq:\*  
47: /cgn2\_6/pdata/1/pna/US6008\_COMB.seq:\*  
48: /cgn2\_6/pdata/1/pna/US6009\_COMB.seq:\*  
49: /cgn2\_6/pdata/1/pna/US6010\_COMB.seq:\*  
50: /cgn2\_6/pdata/1/pna/US6011\_COMB.seq:\*  
51: /cgn2\_6/pdata/1/pna/US6012\_COMB.seq:\*  
52: /cgn2\_6/pdata/1/pna/US6013\_COMB.seq:\*  
53: /cgn2\_6/pdata/1/pna/US6014\_COMB.seq:\*  
54: /cgn2\_6/pdata/1/pna/US6015\_COMB.seq:\*  
55: /cgn2\_6/pdata/1/pna/US6016\_COMB.seq:\*  
56: /cgn2\_6/pdata/1/pna/US6017\_COMB.seq:\*  
57: /cgn2\_6/pdata/1/pna/US6018\_COMB.seq:\*  
58: /cgn2\_6/pdata/1/pna/US6019\_COMB.seq:\*  
59: /cgn2\_6/pdata/1/pna/US6020\_COMB.seq:\*  
60: /cgn2\_6/pdata/1/pna/US6021\_COMB.seq:\*  
61: /cgn2\_6/pdata/1/pna/US6022\_COMB.seq:\*  
62: /cgn2\_6/pdata/1/pna/US6023\_COMB.seq:\*  
63: /cgn2\_6/pdata/1/pna/US6024\_COMB.seq:\*  
64: /cgn2\_6/pdata/1/pna/US6025\_COMB.seq:\*  
65: /cgn2\_6/pdata/1/pna/US6026\_COMB.seq:\*  
66: /cgn2\_6/pdata/1/pna/US6027\_COMB.seq:\*  
67: /cgn2\_6/pdata/1/pna/US6028\_COMB.seq:\*  
68: /cgn2\_6/pdata/1/pna/US6029\_COMB.seq:\*  
69: /cgn2\_6/pdata/1/pna/US6030\_COMB.seq:\*  
70: /cgn2\_6/pdata/1/pna/US6031\_COMB.seq:\*  
71: /cgn2\_6/pdata/1/pna/US6032\_COMB.seq:\*  
72: /cgn2\_6/pdata/1/pna/US6033\_COMB.seq:\*  
73: /cgn2\_6/pdata/1/pna/US6034\_COMB.seq:\*  
74: /cgn2\_6/pdata/1/pna/US6035\_COMB.seq:\*  
75: /cgn2\_6/pdata/1/pna/US6036\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1000	100.0	11463	19	US-09-508-377-10	Sequence 10, Appl
2	43.8	4.4	2000	31	US-09-887-272A-5263	Sequence 5263, Ap
3	41.4	4.1	457	31	US-09-824-536-1968	Sequence 1968, Ap
4	41.4	4.1	457	31	US-09-824-628-1968	Sequence 1988, Ap
5	41.4	4.1	623	27	US-09-698-014-36	Sequence 36, Appl
6	41.4	4.1	79976	31	US-09-534-859-818	Sequence 818, App
7	41.4	4.1	79976	31	US-09-803-736-818	Sequence 818, App
8	41.2	4.1	1059	34	US-09-911-345-3	Sequence 3, Appl1
9	40.6	4.1	82356	27	US-09-692-412-73	Sequence 73, Appl
10	40.6	4.1	82356	31	US-09-803-736-1053	Sequence 1053, Ap
11	40.4	4.0	2085	28	US-09-702-134-29670	Sequence 29670, A
12	40.4	4.0	2085	31	US-09-815-264-74236	Sequence 74236, A
13	40.4	4.0	2140	24	US-09-620-392-1096	Sequence 1096, Ap
14	40.4	4.0	3554	24	US-09-702-134-92163	Sequence 9663, Ap
15	40.4	4.0	3554	31	US-09-815-264-86646	Sequence 86646, A
16	40.4	4.0	3808	24	US-09-620-392-45673	Sequence 45673, A
17	40.4	4.0	10079	24	US-09-620-392-26402	Sequence 26402, A
18	40.4	4.0	10079	28	US-09-702-134-24210	Sequence 24210, A
19	40.4	4.0	10079	31	US-09-815-264-67802	Sequence 67802, A
20	39.2	3.9	7218	8	US-08-466-194-14	Sequence 14, Appl
21	39.2	3.9	1055	31	US-09-806-708A-23	Sequence 23, Appl
22	39.2	3.9	32768	60	US-06-212-656-26	Sequence 26, Appl
23	39.2	3.9	832	24	US-09-621-976-2813	Sequence 2813, Ap
24	39.2	3.9	832	53	US-06-147-499-2813	Sequence 2813, Ap
25	39.2	3.9	1169	56	US-06-172-377-2283	Sequence 2283, Ap
26	38.8	3.9	17168	28	US-09-702-134-25673	Sequence 25673, A
27	38.8	3.9	17168	31	US-09-815-264-73103	Sequence 73103, A
28	38.8	3.9	17201	24	US-09-620-392-4351	Sequence 4351, Ap
29	38.2	3.8	2000	33	US-09-887-272A-5263	Sequence 5263, Ap
30	37.8	3.8	712	61	US-06-226-326-6028	Sequence 6028, Ap
31	37.8	3.8	32768	59	US-60-208-129-68	Sequence 68, Appl

## ALIGNMENTS

```

RESULT 1
US-09-508-377-10
: Sequence 10, Application US/09508377
: GENERAL INFORMATION:
: APPLICANT: KALEEN, ZHONGYILI
: APPLICANT: MORELL, MATTHEW
: APPLICANT: RAHMAN, SADEOUR
: TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS
: FILE REFERENCE: 054270/0126
: CURRENT APPLICATION NUMBER: US/09/508,377
: CURRENT FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: AU Pp 2509
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: PCT/AU98/00743
: PRIOR FILING DATE: 1998-09-11
: PRIOR APPLICATION NUMBER: AU Pp 9108
: PRIOR FILING DATE: 1997-09-12
: NUMBER OF SEQ ID NOS: 71
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 11463
: TYPE: DNA
: ORGANISM: Trifolium tauschii
US-09-508-377-10

```

Query Match	100.0%;	Score 1000;	DB 19;	Length 11463;
Best Local Similarity	100.0%;	Pred. No. 2.3e-282;		
Matches 1000;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	agaaacacccatcttagatattttttttgtctcttttcgagcgtgagtcgctggagaga	60
Db	1	agaaacacccatcttagatattttttttgtctcttttcgagcgtgagtcgctggagaga	60
QY	61	tttagcgtctagttttcttaaaagaaacagccattagccctctttacaaaagctcaa	120
Db	61	tttagcgtctagttttcttaaaagaaacagccattagccctctttacaaaagctcaa	120
QY	121	ccagttccaaaacgctctcttagatcacccagctgtaaaagttaagcggagaccccaac	180
Db	121	ccagttccaaaacgctctcttagatcacccagctgtaaaagttaagcggagaccccaac	180
QY	181	aggcgcatctcgaaacttgacagacagctcacagagagacccagcaacaaagcttgagctg	240
Db	181	aggcgcatctcgaaacttgacagacagctcacagagagacccagcaacaaagcttgagctg	240
QY	241	acaaacgagcgtgagtcggtgacaaatgggttcattctatggcggtcggagcaagaaagaa	300
Db	241	acaaacgagcgtgagtcggtgacaaatgggttcattctatggcggtcggagcaagaaagaa	300
QY	301	gacgcacatgaaacccaatgatatgctataagccctatgagaggaacaaacctgtgaact	360
Db	301	gacgcacatgaaacccaatgatatgctataagccctatgagaggaacaaacctgtgaact	360

QY	361	ttcccccttggaattatbtagtcgcacaactttttttaaitygaagcaaggttggcaac	420
Db	361	ttcccccttggaattatbtagtcgcacaactttttttaaitygaagcaaggttggcaac	420
QY	421	acatcattttcaacaaggaataattatctcaaacacacatgacatgcattctcaaa	480
Db	421	acatgattttcaacaaggaataattatctcaaacacacatgacatgcattctcaaa	480
QY	481	ccatgaccgcagtgcccatgvcgaggtvggaaacgaagaacttgaaatcaacatccagtt	540
Db	481	ccatgaccgcagtcgcatgvcgaggtvggaaacgaagaacttgaaatcaacatccagtt	540
QY	541	gtcgcgtctgaaagaaggtgtgacactgtgaaagatgctgattacagatttcatttcatacat	600
Db	541	gtcgcgtctgaaagaaggtgtgacactgtgaaagatgctgattacagatttcatttcatacat	600
QY	601	gtacacatacataatgtagcccttacaattgttttttggagcagaagtgtgtgtctttt	660
Db	601	gtacacatacataatgtagcccttacaattgttttttggagcagaagtgtgtgtctttt	660
QY	661	ttttaacacgaataatgcatatgctcgtgcccgcatagtcgtgcagatcgtgatcgtgcgag	720
Db	661	ttttaacacgaataatgcatatgctcgtgcccgcatagtcgtgcagatcgtgatcgtgcgag	720
QY	721	acgagcgacaatcgcagacactaccacactgctttgttctggagacacaataattttttgt	780
Db	721	acgagcgacaatcgcagacactaccacactgctttgttctggagacacaataattttttgt	780
QY	781	aaacacaataataactctttaaacgaggttactagaagccgctaacgcatcgtgcaggttaa	840
Db	781	aaacacaataataactctttaaacgaggttactagaagccgctaacgcatcgtgcaggttaa	840
QY	841	acgcgtctccagccgttgtgttgcgattcgtctcctccgcagcagcgtcgtcctccacg	900
Db	841	acgcgtctccagccgttgtgttgcgattcgtctcctccgcagcagcgtcgtcctccacg	900
QY	901	tcgcgcgtctgcctgcacactctgtgtgcgcgcgcaagaaggagaggaagcaagccgcg	960
Db	901	tcgcgcgtctgcctgcacactctgtgtgcgcgcgcaagaaggagaggaagcaagccgcg	960
QY	961	cacacacactcacacacgcgcacatctcccgctggtgtccct 1000	
Db	961	cacacacactcacacacgcgcacatctcccgctggtgtccct 1000	
RESULT 2			
US-09-887-272A-5263/C			
Sequence 5263, Application US/09887272A			
GENERAL INFORMATION:			
APPLICANT: Hou, Yu-Ming			
APPLICANT: Quan, Sheng			
APPLICANT: Chang, Hur-Song			
APPLICANT: Zhu, Tong			
APPLICANT: Whitlam, Steve			
APPLICANT: Goff, Steve			
APPLICANT: Glazedbrook, Jane			
APPLICANT: Chen, Wenqiong			
APPLICANT: Katagiri, Fumiaki			
APPLICANT: Xie, Zhiyi			
APPLICANT: Tao, Yi			
APPLICANT: Zou, Guangzhou			
APPLICANT: Cooper, Bret			
TITLE OF INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST			
FILE REFERENCE: 1360.0030S2			
CURRENT APPLICATION NUMBER: 2001-06-23			
CURRENT FILING DATE: 2001-06-23			
PRIOR APPLICATION NUMBER: 60/213,634			
PRIOR FILING DATE: 2000-06-23			
PRIOR APPLICATION NUMBER: 60/214,926			
PRIOR FILING DATE: 2000-06-23			
PRIOR APPLICATION NUMBER: 60/261,320			
PRIOR FILING DATE: 2001-01-12			

```

; PRIOR APPLICATION NUMBER: 60/264,353
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/273,879
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 09/887,271
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 6813
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 5263
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-887-272A-5263

```

```

Query Match      4.4%; Score 43.8; DB 33; Length 2000;
Best Local Similarity 10.2%; Pred. No. 0.37;
Matches 101; Conservative 426; Mismatches 458; Indels 8; Gaps 3;

```

```

QY 8 ccccatcttagatctttttttgtctctcttgagcggtgctgagagatagcgt 67
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 995 YMAISSTMKWYAWKRAVYAMSRSKRYTWCTGKMAITCGTKMAAGKWRMAMCWT 936
QY 68 ctatgtttctaaagaacagcgccattagccctgcttaacaaaggctcaaccagctc 127
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 935 CCMWKKMKMTSCMMKWTWCTWCTWCTWCTWCTWCTWCTWCTWCTWCTWCTWCTW 876
QY 128 aaaagctctgtagatcaccagctgcaaaagtaagcgagagaccacaaacagcgca 187
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 875 AWTWMTCMCAKMAKWTGATWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTW 816
QY 188 ttgcgaactgagacagcgctcagcagagagagagagagagagagagagagagag 247
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 815 TSWRMAAMAGMKRKRKRAVYAMRWRMCKAGMARMKSRIRMKWKATATYMKMA 757
QY 248 acgtgagtcgtgacacatgagtgatctatgagcgctcgagagagagagagagagac 307
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 756 MTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTWMTW 697
QY 308 atgaacacatgagtgatctatgagcgctgagagagagagagagagagagagagag 367
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 696 YMKRKMWTYRYRYRSCMTYRAMSKRKWKAGASMKSCWMTWMTWMTWMTWMTWMTW 637
QY 368 ttgcgaactgagacagcgctcagcagagagagagagagagagagagagagagag 427
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 636 KTRRYTSSYMTGWTGWTSSYMSMTSKMSMTGKMTCTWYTSKMSGSTRSRMGRSGM 577
QY 428 ttctcaacaaagaaataatctcaaacacacatgacatgacatctcaaacatgac 487
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 576 SRMYMMMKMKMRKRYMYRMKMKCTWRRCMYRMGYTMTYTSRMYMTGRTYARVTSKR 517
QY 488 ccgacgagtcacatgagtgagagagagagagagagagagagagagagagagagag 547
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 516 RYMYMYRYRKYCMY----YYGYMYKCSYMYRKYCKACCKCCYACMKWKAAGMMYMY 462
QY 548 cgaagagagagatgacatgagagagagagagagagagagagagagagagagagagag 607
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 461 RKYSKMMRSTKYMMSMYKCKSMKYGAKGCGCKMWTYCSGYMKWYMYGSKYSKRC 402
QY 608 tacataatgacccataatgttttttggagagagagagagagagagagagagagagag 667
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 401 YKYMRYMYKGMWYMYYSMSMTWYYYAKYKMYKMYKMYKMYKMYKMYKMYKMYK 342
QY 668 acgaaatgacatgagtgagagagagagagagagagagagagagagagagagagag 727
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 341 CYMKCKMYRMRMKRYKYSRCYCMWATYWCOCYKRRKYSRMSMRMAGKMKMR 282
QY 728 acatacagacacacacacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 787
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 281 SWGRMCRSYMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYKMYK 222
QY 788 ataataactataacgagagagagagagagagagagagagagagagagagagagagag 845
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

```

```

DB 221 RMYSACRRYSRTSYCGSCYCGSSKMKMYMSKSCSMRMTSCSCCYTCYAGMCSGM 162
QY 846 ctcccacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 905
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 161 SMYMSGCGYTRGMRKRSYMSCKRKYSCCTKRYCSYGYRYCKWYKYSKYCYC 102
QY 906 ccgtgctgcccctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 965
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 101 YWYMTYMYKMKCSRSSMSWCAVCTSTSRMMSMTYAAKMKMGSSGMYRMSKSC 42
QY 966 acatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 98
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 41 KMYSKYSSCKYTGSKCKTKRKYCYWSSGYSM 9

```

```

RESULT 3
US-09-824-556-1988
; Sequence 1988, Application US/09824556
; GENERAL INFORMATION:
; APPLICANT: Hunter, John J.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2084-001
; CURRENT APPLICATION NUMBER: US/09/824,556
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/193,423
; NUMBER OF SEQ ID NOS: 9973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1988
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(457)
; OTHER INFORMATION: n = A,T,C or G
US-09-824-556-1988

```

```

Query Match      4.1%; Score 41.4; DB 31; Length 457;
Best Local Similarity 43.4%; Pred. No. 0.85;
Matches 75; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```

```

QY 627 ttgttttttggagagagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 686
   ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 136 ttcttttttaaaaaaaaggggggggttttttttttttttttttttttttttttttt 195
QY 687 ccgcatgctgtagatcgtagatcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 746
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 196 tttnnccctlnnccccccngngngngggnnnnnnnnnnnnnnnnnnnnnnnnnnn 255
QY 747 ctgctttgtctgggacacataatgtttttttaaacaataataactat 799
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 256 annntttttttnncccccgtaaaacctlnnnccccaaaaaaataaat 308

```

```

RESULT 4
US-09-824-628-1988
; Sequence 1988, Application US/09824628
; GENERAL INFORMATION:
; APPLICANT: Hunter, John J.
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2084-001
; CURRENT APPLICATION NUMBER: US/09/824,628
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/193,423
; NUMBER OF SEQ ID NOS: 9973
; SOFTWARE: FastSeq for Windows Version 4.0

```

```
; SEQ ID NO 1988
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(457)
; OTHER INFORMATION: n = A,T,C or G
US-09-824-628-1988
```

```
Query Match
Best Local Similarity 4.1%; Score 41.4; DB 31; Length 457;
Best Local Similarity 43.4%; Pred. NO. 0.85;
Matches 75; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
```

```
OY 627 ttgttttttgagcagagtggtgtgtgttttttttttaaccggaataatgcacatgctg 686
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 136 ttttttttaaaaaaaagggggggggttttttttttttttttttttttttttttttt 195
OY 687 ccgcatgctgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcat 746
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 196 tttnccctttnnccccccnngnngnngnngnngnngnngnngnngnngnngnngn 255
OY 747 ctgctttgtctggagacacataatgtttttttaacacaaataacttat 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 annntttttttnnccccccgtaaaaacctttnnnccaaaaaaataanaat 308
```

```
RESULT 5
US-09-698-014-36
; Sequence 36, Application US/09698014
; GENERAL INFORMATION:
; APPLICANT: Gearling, David P.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2014-001
; CURRENT APPLICATION NUMBER: US/09/698,014
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/162,363
; NUMBER OF SEQ ID NOS: 6098
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(623)
; OTHER INFORMATION: n = A,T,C or G
US-09-698-014-36
```

```
Query Match
Best Local Similarity 4.1%; Score 41.4; DB 27; Length 623;
Best Local Similarity 50.2%; Pred. NO. 1;
Matches 102; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```
OY 336 cactcttccctctggaatcatagctcacacttttttttaatggaagaagtgtg 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 374 cccatttccctctctcacttcttctcagcagttagctaattcaagatcgagaatg 433
OY 416 caaacatgcatcttcaacaagaataatattctcaaccacatgacatgcatc 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 434 gttacagaacttttcaacaagaagttacatagcagaatgctccagaattga 493
OY 476 tcaaccatgcccgcagcagtcacgagtgagtggaacgaagaactggaataacacatcc 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 494 tcagataagagcccgcatgttgtagattttgggaagaagaagactgagttctattcc 553
OY 536 cagttgctgagtcgagaagaga 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
DB 554 cacttctcaattgggagaagga 576
```

```
RESULT 6
US-09-534-859-818
; Sequence 818, Application US/09534859
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Last, Robert L.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
; FILE REFERENCE: 38-10(15493)B
; CURRENT APPLICATION NUMBER: US/09/534,859
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 1127
; SEQ ID NO 818
; LENGTH: 79976
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-534-859-818
```

```
Query Match
Best Local Similarity 4.1%; Score 41.4; DB 20; Length 79976;
Best Local Similarity 47.8%; Pred. NO. 14;
Matches 120; Conservative 0; Mismatches 131; Indels 0; Gaps 0;
```

```
OY 427 atttcaacaagaagaataattctcaaacaccatgacatgacatctccaaccatgc 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 40914 attaatctctgagaatttggtgtatgtctcctcaaatatttaataatttagcacac 40973
OY 487 accgcagatgcacatgcaggtggaacgaagaactgaaataacatcccgattgcag 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 40974 atagacaactatttaagaatacaacataaattctgtcgtatatttattgtttaat 41033
OY 547 tggagaagagtgacacacgaagatgctgtatgattcattatctacatgataca 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41034 tagagggcaaaaaggcaagaatcttattcttttaaatgtatttttaggaagaacaat 41093
OY 607 atacatagttaccctacacattgttttttgagagagagtggtgtctttttttta 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41094 atataatgactaataataaagtgtttttgtgtcctatctgattgagattattatgt 41153
OY 667 cagcaaatgc 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 41154 gaatcaattc 41164
```

```
RESULT 7
US-09-803-736-818
; Sequence 818, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; PRIOR FILING DATE: 2000-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 818
; LENGTH: 79976
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```

```
US-09-803-736-818
; Sequence 818, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; PRIOR FILING DATE: 2000-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 818
; LENGTH: 79976
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```



Query Match	Similarity	4.1%	Score 40.6	DB 31	Length 82356
Best Local	Similarity	52.7%	Pred. No. 25		
Matches	88	Conservative	0	Mismatches	79
				Indels	0
				Gaps	0
368	tcctggaatctcagctcacacctttttttaaagtgaagcaagagttgycacaacatcyca	427			
43442	tcctcaactagttacctcttaagtgttttttttaataaaaaaacgaatccaatccatcaatgcga	43863			

```

QY 428 ttccaacaagaataattatctcaaacaccatgcacatctccaacatgca 487
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43382 TCAAACCAAAATTAATTAATTCAATTAATGATCAATTTTGTATTTGAAAAAAT 43323
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 488 ccgacagatccatgcagagtggaacgaagactgaataacacatc 534
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43322 TTGAATAGTCAGAACCAATTAACCGAATTAACGATATATCCACTTC 43276
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 11

```

US-09-702-134-29670/c
; Sequence 29670, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 29670
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-702-134-29670

```

```

Query Match          4.0%; Score 40.4; DB 28; Length 2085;
Best Local Similarity 50.5%; Pred. No. 3.8;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```

```

QY 765 caataaagtgttctgtaacaaataactataacgaaggtactagaagcgctaa 824
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1530 CATCATTTCTGATTGCGAGTTAAGTTAATTTGACAAAAAGAGCATGAAGCCTTAGA 1471
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 cggcatgcccaggtaaacgcgcgtccagcgggtgttgatcgtcgtcccgcaagc 884
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1470 AGAAGAAAAAGAAAAAGTTAAGATTGGGGGGGCGCTGCTGCTTCTGACGGCC 1411
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 agcgtcctccacgcgtccgcgtccacccctcgtcgtgcgcgcgaagaagga 944
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1410 AGGTCCCGCGGCTGTTGCCACCGTGCCTTGCACTGCTGCCCTCCCTCTGTCGGAG 1351
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 945 ggaagaacgaagc 958
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1350 TCGGGGGCGCAAGC 1337
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 12

```

US-09-815-264-74236/c
; Sequence 74236, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31

```

```

; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 74236
; LENGTH: 2085
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-74236

```

```

Query Match          4.0%; Score 40.4; DB 31; Length 2085;
Best Local Similarity 50.5%; Pred. No. 3.8;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```

```

QY 765 caataaagtgttctgtaacaaataactataacgaaggtactagaagcgctaa 824
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1530 CATCATTTCTGATTGCGAGTTAAGTTAATTTGACAAAAAGAGCATGAAGCCTTAGA 1471
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 cggcatgcccaggtaaacgcgcgtccagcgggtgttgatcgtcgtcccgcaagc 884
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1470 AGAAGAAAAAGAAAAAGTTAAGATTGGGGGGGCGCTGCTGCTTCTGACGGCC 1411
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 agcgtcctccacgcgtccgcgtccacccctcgtcgtgcgcgcgaagaagga 944
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1410 AGGTCCCGCGGCTGTTGCCACCGTGCCTTGCACTGCTGCCCTCCCTCTGTCGGAG 1351
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 945 ggaagaacgaagc 958
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1350 TCGGGGGCGCAAGC 1337
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 13

```

US-09-620-392-1096/c
; Sequence 1096, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 1096
; LENGTH: 2140
; TYPE: DNA
; ORGANISM: Oryza sativa
; OTHER INFORMATION: unsure at all n locations
US-09-620-392-1096

```

```

Query Match          4.0%; Score 40.4; DB 24; Length 2140;
Best Local Similarity 50.5%; Pred. No. 3.9;
Matches 98; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```

```

QY 765 caataaagtgttctgtaacaaataactataacgaaggtactagaagcgctaa 824
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1585 CATCATTTCTGATTGCGAGTTAAGTTAATTTGACAAAAAGAGCATGAAGCCTTAGA 1526
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 cggcatgcccaggtaaacgcgcgtccagcgggtgttgatcgtcgtcccgcaagc 884
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1525 AGAAGAAAAAGAAAAAGTTAAGATTGGGGGGGCGCTGCTGCTTCTGACGGCC 1466
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 agcgtcctccacgcgtccgcgtccacccctcgtcgtgcgcgcgaagaagga 944
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1465 AGGTCCCGCGGCTGTTGCCACCGTGCCTTGCACTGCTGCCCTCCCTCTGTCGGAG 1406
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 945 ggaagaacgaagc 958
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1405 TCGGGGGCGCAAGC 1392
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 14

```

US-09-702-134-9663

```

```

: Sequence 9663, Application US/09702134
: GENERAL INFORMATION:
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Gao, Yongwei
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jindong
: APPLICANT: Mcininch, James
: APPLICANT: Wu, Wei
: TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
: FILE REFERENCE: 38-21(51237)F
: CURRENT APPLICATION NUMBER: US/09/702,134
: CURRENT FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 52202
: SEQ ID NO 9663
: LENGTH: 3554
: TYPE: DNA
: ORGANISM: Oryza sativa
: US-09-702-134-9663

```

Query Match	4.0%;	Score 40.4;	DB 28;	Length 3554;
Best Local Similarity	50.5%;	Pred. No. 5.1;		
Matches 98;	Conservative 0;	Mismatches 96;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 15
US-09-815-264--86646
? Sequence 86646, Application US/09815264
? GENERAL INFORMATION:
? APPLICANT: Boukharov, Andrey A.
? APPLICANT: Cao, Yongwei
? APPLICANT: Dotson, Stanton B.
? APPLICANT: Koshi, Jeffrey M.
? APPLICANT: Kovalic, David K.
? APPLICANT: Liu, Jingtong
? APPLICANT: McIninch, James
? APPLICANT: Wu, Wei
? TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
? FILE REFERENCE: 38-21(51237)G
? CURRENT APPLICATION NUMBER: US/09/815,264
? CURRENT FILING DATE: 2001-03-23
? PRIOR APPLICATION NUMBER: US 09/520,392
? PRIOR FILING DATE: 2000-07-19
? PRIOR APPLICATION NUMBER: US 09/702,134
? PRIOR FILING DATE: 2000-10-31
? NUMBER OF SEQ ID NOS: 109669
? SEQ ID NO 86646
? LENGTH: 3554
? TYPE: DNA
? ORGANISM: Oryza sativa
US-09-815-264--86646

```

Query Match	4.0%;	Score 40.4;	DB 31;	Length 3554;
Best Local Similarity	50.5%;	Pred. No. 5.1;		
Matches 98;	Conservative 0;	Mismatches 96;	Indels 0;	Gaps 0;

OY	765	caactaaatgttttttgcatacaaaaataaactatcataaagagagcactaaagggccgcctaa	824
Db	1119	catcatcttcgatctgcgagatlaaagttcaatttgcacaaaaaagagcactaagaagccctaga	1178
OY	825	cgcgcattgcacgaagaaacgcgcgtccccaacccgttgcgttgcatactcgcctcccgcaagc	884
Db	1179	agaaagaagaaaaaagaagtaagatgtgcgcgggggcgttcggtccattctgaacgcgcgc	1238
OY	885	agagtcgcgtccacacgcgtccgctgcgtgcgcacccctcgtctgtgcgcgcgcagaaga	944
Db	1239	agctcgcgcgcgtccgttgcgcacacgcctgcgttcgcactcgtgcgttcccccctcgtcgcag	1298
OY	945	ggaagaacgcgaacgc	958
Db	1299	tcggggggcgcaacgc	1312

Search completed: September 12, 2002, 10:08:12  
Job time: 9440 sec

0 20

GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 06:06:43 : Search time 643.61 Seconds  
(without alignments)  
4916.495 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_1\_1000

Perfect score: 1000

Sequence: 1 agaacacccctcatttaga.....acactcccggtggtccct 1000

Scoring table:

IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2084544 seqs, 1582152599 residues

Total number of hits satisfying chosen parameters: 4169088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing First 45 summaries

Database :

Pending\_Patents\_NA\_New:\*  
1: /cgn2\_6/pdata/2/pna/PCT\_US06\_NEW\_COMB.seq:\*  
2: /cgn2\_6/pdata/2/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/pdata/2/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/pdata/2/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/pdata/2/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/pdata/2/pna/US10\_NEW\_COMB.seq:\*  
7: /cgn2\_6/pdata/2/pna/US10\_NEW\_COMB.seq:\*  
8: /cgn2\_6/pdata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37.4	3.7	1191139	US-09-811-352B-1	Sequence 1, Appl
2	37.4	3.7	1191139	US-10-140-924-1	Sequence 1, Appl
3	37.4	3.7	1691139	US-10-067-514-1	Sequence 10980, A
4	37.2	3.7	6493	PCT-US02-25766-10980	Sequence 10980, A
5	35.6	3.6	691	US-10-155-881-25269	Sequence 25269, A
6	35.6	3.6	1275	US-10-027-632-213914	Sequence 213914, A
7	35.6	3.6	1741	US-10-201-386-12	Sequence 12, Appl
8	35.4	3.5	2407	US-10-198-846-13552	Sequence 13552, A
9	35.2	3.5	774	US-10-155-881-23995	Sequence 23995, A
10	35	3.5	1754	US-10-155-881-12490	Sequence 12490, A
11	35	3.5	2991	US-10-179-131-2533	Sequence 2533, Ap
12	34.4	3.4	10495	US-09-053-375B-600	Sequence 600, App
13	34.4	3.4	15932	PCT-US02-09105-482	Sequence 482, App
14	34.4	3.4	15932	PCT-US02-09370-1049	Sequence 1049, Ap
15	34.4	3.4	15932	US-10-105-299-6415	Sequence 6415, Ap
16	34.2	3.4	1449	US-10-212-760-275	Sequence 275, App
17	34.2	3.4	1792	US-10-027-632-250431	Sequence 250431, A
18	34	3.4	3422	US-09-979-715-13	Sequence 12, Appl
19	34	3.4	5791	US-09-979-715-13	Sequence 13, Appl
20	33.6	3.4	218336	PCT-US02-11086-10	Sequence 10, Appl
21	33.4	3.3	299	US-09-539-331D-20130	Sequence 20130, A
22	33.4	3.3	497	US-09-785-276A-60792	Sequence 60792, A
23	33.4	3.3	1292	US-09-634-754C-493	Sequence 493, App
24	33.4	3.3	1292	US-09-634-754D-493	Sequence 493, App
25	33.4	3.3	2324	US-09-629-469A-14737	Sequence 14737, A

26	33.2	3.3	168575	US-10-178-194-1	Sequence 1, Appl
27	33	3.3	495	US-09-620-393B-1200	Sequence 1200, Ap
28	33	3.3	495	US-09-935-625-6643	Sequence 6643, Ap
29	33	3.3	495	US-09-935-625-18919	Sequence 18919, A
30	33	3.3	567	US-10-027-632-43342	Sequence 43342, A
31	33	3.3	567	US-10-027-632-43342	Sequence 43342, A
32	33	3.3	567	US-10-027-632-66444	Sequence 66444, A
33	33	3.3	567	US-10-027-632-66445	Sequence 66445, A
34	33	3.3	567	US-10-027-632-311165	Sequence 311165, A
35	33	3.3	567	US-10-027-632-311166	Sequence 311166, A
36	33	3.3	854	US-10-027-632-151631	Sequence 151631, A
37	33	3.3	992	US-10-144-860-111	Sequence 111, App
38	33	3.3	2497	US-09-705-256A-112	Sequence 112, App
39	32.8	3.3	166	US-10-113-872-1580	Sequence 1580, Ap
40	32.8	3.3	517	US-10-027-632-281359	Sequence 281359, Ap
41	32.8	3.3	1859	US-09-705-256A-6386	Sequence 6386, Ap
42	32.8	3.3	4649	US-09-919-002-1530	Sequence 1530, Ap
43	32.8	3.3	13578	US-10-105-299-11948	Sequence 11948, A
44	32.6	3.3	459	US-10-098-754-659	Sequence 659, App
45	32.6	3.3	507	US-10-027-632-194618	Sequence 194618, A

#### ALIGNMENTS

RESULT 1  
US-09-811-352B-1  
; Sequence 1, Application US/09811352B  
; GENERAL INFORMATION:  
; APPLICANT: Jonsdottir, Solveig  
; APPLICANT: Grelarsdottir, Solveig  
; APPLICANT: Jonsdottir, Sif  
; APPLICANT: Reynisdottir, Sigridur Th.  
; TITLE OF INVENTION: HUMAN STROKE GENE  
; FILE REFERENCE: 2345.2010-000  
; CURRENT FILING DATE: 2002-05-03  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1191139  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-811-352B-1

Query Match 3.7%; Score 37.4; DB 5; Length 1191139;  
Best Local Similarity 52.2%; Pred. No. 63;  
Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 390 ttttttaaggaagcaaggttgcaaacacatgcatltaaacagaagaaat 449  
Db 287753 ttatgattactagacacaggactattttatcaactatcttaacatacaaaat 287812

QY 450 tccaacaccacatgacatctcaaacacatgaccacagcagtcagagtg 509  
Db 287813 tgcataattcttggaatcatcatccacttaaatataggaagaaat 287872

QY 510 aacagaagactgaataacatccacatgctgcagtc 548  
Db 287873 atataacatgaataatccacatccagctgaagaaatc 287911

RESULT 2  
US-10-140-924-1  
; Sequence 1, Application US/10140924  
; GENERAL INFORMATION:  
; APPLICANT: Grelarsdottir, Solveig  
; APPLICANT: Jonsdottir, Sif  
; APPLICANT: Reynisdottir, Sigridur Th.  
; TITLE OF INVENTION: HUMAN STROKE GENE  
; FILE REFERENCE: 2345.2010-000  
; CURRENT FILING DATE: 2002-05-07

```
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1191139
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-924-1
```

```
Query Match          3.7%; Score 37.4; DB 7; Length 1191139;
Best Local Similarity 52.2%; Pred. No. 63;
Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
OY 390 tttttttaaaggaagagttggcaacacatgcatctttcaacaagaataat 449
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 287753 ttatgattactgcaacagggactttttatccactactcttaacataaattat 287812

OY 450 tctcaaacacatgacatgcaatctcacaaccatgacggagatgcagagtg 509
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 287813 tgtaaatttctaggaattcatatccacttaaaatagagaagatcatttggaaaaa 287872

OY 510 aacgaagactgaaatcaacatccacatgtgtcgagtc 548
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 287873 atattaacataagaaatcagcatcccgattagaatc 287911
```

```
RESULT 3
US-10-067-514-1
; Sequence 1, Application US/10067514
; GENERAL INFORMATION:
; APPLICANT: GretaSodttr, Solveig
; APPLICANT: Jonsdottir, Sif
; APPLICANT: Reyndsdottir, Sigridur Th.
; TITLE OF INVENTION: HUMAN STROKE GENE
; FILE REFERENCE: 2345, 2010-003
; CURRENT APPLICATION NUMBER: US/10/067,514
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 09/811/352
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1691139
; TYPE: DNA
; ORGANISM: Human
US-10-067-514-1
```

```
Query Match          3.7%; Score 37.4; DB 7; Length 1691139;
Best Local Similarity 52.2%; Pred. No. 71;
Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
OY 390 tttttttaaaggaagagttggcaacacatgcatctttcaacaagaataat 449
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 787752 ttatgattactgcaacagggactttttatccactactcttaacataaattat 787811

OY 450 tctcaaacacatgacatgcaatctcacaaccatgacggagatgcagagtg 509
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 787812 tgtaaatttctaggaattcatatccacttaaaatagagaagatcatttggaaaaa 787871

OY 510 aacgaagactgaaatcaacatccacatgtgtcgagtc 548
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 787872 atattaacataagaaatcagcatcccgattagaatc 787910
```

```
RESULT 4
PCT-US02-25766-10980/c
; Sequence 10980, Application PC/TUS0225766
; GENERAL INFORMATION:
; APPLICANT: GENE LOGIC, INC.
; APPLICANT: MUNGER, William E
; APPLICANT: PAULK, Ronald
; APPLICANT: SUN, Hongwei
```

```
; APPLICANT: SASAI, Hitoshi
; APPLICANT: WAGA, Iwao
; APPLICANT: YAMAMOTO, Jun
; TITLE OF INVENTION: Gene Expression Profiles in Glomerular Diseases
; FILE REFERENCE: 44921-5068-No
; CURRENT APPLICATION NUMBER: PCT/US02/25766
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/311,837
; PRIOR FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 13946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10980
; LENGTH: 6493
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. X98534
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(6493)
; OTHER INFORMATION: n = a or c or g or t
PCT-US02-25766-10980
```

```
Query Match          3.7%; Score 37.2; DB 1; Length 6493;
Best Local Similarity 52.6%; Pred. No. 9.4;
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
```

```
OY 598 catgacaatacatcagaatgtaaccctacaattgttttggcagagatggtgctt 657
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3453 CAAGTCCAGTATTTTGGGGCCCAAGTATTTTGGTTTGGTTTGTGTGTGT 3394

OY 658 ttttttcaagaatgcatagctggcccgatgctgctgagatcgatcgatcgctg 717
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3393 TTGTTTGAAGACATGACATTTCTATTCTCAGGCTGGAGTGCAGCATGATGCTCG 3334

OY 718 gagaagcggagcaatcagacactcaacactgct 751
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3333 CTGAGGCTCACTCTCGGCTCAAGTATCTCT 3300
```

```
RESULT 5
US-10-155-881-25269/c
; Sequence 25269, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kowalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Luthiya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 25269
; LENGTH: 691
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-155-881-25269
```

```
Query Match          3.6%; Score 35.6; DB 7; Length 691;
Best Local Similarity 54.6%; Pred. No. 11;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
OY 828 catgcaagataaagcgtcccaagcgtgtgttgcgactcgtcctccgcagcagc 887
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 173 CCGTGCATGTGGCCCTCTCGCGCACTTGTAGACAGCGCGCTGCTCCGCCGCCGCC 114

OY 888 gtgcctccacgtctgcgtcgtcgtccacctgtgtgtgctgcgcgcagcagaaggag 947
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



Query Match	3.5%	Score 35.4;	DB 6;	Length 2407;
Best Local Similarity	48.3%	Pred. No. 21;		
Matches 99;	Conservative 0;	Mismatches 106;	Indels 0;	Gaps 0

[illegible]

## RESULT 9

US-10-155-881-23995/c  
; Sequence 23995, Application US/10155881

```

: GENERAL INFORMATION:
: APPLICANT: Dotson, Stanton B.
: APPLICANT: Kowalic, David K.
: APPLICANT: Liu, Jindong
: APPLICANT: Lutfiyya, Linda L.
: APPLICANT: McIndinch, James
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
: TITLE OF INVENTION: TRANSCRIPTION IN PLANTS
: FILE REFERENCE: 38-21(15300)I
: CURRENT APPLICATION NUMBER: US/10/155,881
: CURRENT FILING DATE: 2002-05-22
: NUMBER OF SEQ ID NOS: 37595
: SEQ ID NO 23995
: LENGTH: 774
: TYPE: DNA
: ORGANISM: Oryza sativa
: OS-10-155-881-23995

```

Query Match	3.5%;	Score 35.2;	DB 7;	length 774;
Best Local Similarity	60.4%;	Pred. No. 15;		
Matches	58;	Conservative	0;	Mismatches 38;
			Indels	0;
			Gaps	0;

[illegible]

RESULT 10  
HS-10-155-

Sequence 12490, Application US/10155881

```

: GENERAL INFORMATION:
: APPLICANT: Dotson, Stanton B.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: Lutfiyya, Linda L.
: APPLICANT: McIninch, James
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
: TITLE OF INVENTION: TRANSCRIPTION IN PLANTS
: FILE REFERENCE: 38-21(15300)J
: CURRENT APPLICATION NUMBER: US/10/155,881
: CURRENT FILING DATE: 2002-05-22
: NUMBER OF SEQ ID NOS: 37595
: SEQ ID NO 12490

```

```

; LENGTH: 1754
; TYPE: DNA
; ORGANISM: Glycine max
US-10-155-881-12490

```

Query Match	3.5%;	Score 35;	DB 7;	Length 1754;
Best Local Similarity	50.9%;	Pred. NO. 24;		
Matches	83;	Conservative	0;	Mismatches 80;
			Indels	0;
			Gaps	0;

Oy	551	gaagagagatgcactgaaagratgcygtrattacgatttcatttacatcatgtaacaatac	610
Db	522	GATGAGATGTATGAGCAAAAACCTCGGAGTATTAATACATATATMAAATATGGAAGATTGG	463
Oy	611	ataagttacctacaatttgttttttggagagagtggtgtgtgtctttttttttaaag	670
Db	462	ATTATGTAATTTGTATTTCATMAAATGAAAAGCGTATTTTTTTTTTTTTTTTTTTTAAAG	403
Oy	671	aaatgcatagctggcccgagatcgctgcagaatcgagatgcg	713
Db	402	TAAATGTCTACGTGTGAGGAGATTAAAGGACAGAGTAGCTTG	360

## RESULT 11

US-10-179-131-2533/c  
; Sequence 2533, Application US/10179131

```

1  GENERAL INFORMATION:
2  APPLICANT: HARE, ROBERTA S.
3  APPLICANT: SHAW, KAREN J.
4  APPLICANT: SHIMER JR., GEORGE H.
5  APPLICANT: KESSLER, MARCO
6  APPLICANT: NOLLING, JORK
7  APPLICANT: ZENG, QIANDONG
8  APPLICANT: GREENE, JONATHAN R.
9  TITLE OF INVENTION: CANDIDA ALBICANS NUCLEIC ACIDS AND POLYPEPTIDES
10 TITLE OF INVENTION: AND USES THEREFOR
11 FILE REFERENCE: 2976-4031
12 CURRENT APPLICATION NUMBER: US/10/179,131
13 CURRENT FILING DATE: 2002-06-21
14 NUMBER OF SEQ ID NOS: 10194
15 SEQ ID NO 2533
16 LENGTH: 2991
17 TYPE: DNA
18 ORGANISM: Candida albicans
19 US-10-179-131-2533

```

Query Match	3.5%	Score 35;	DB 7;	Length 2991;
Best Local Similarity	48.7%;	Pred. No. 29;		
Matches 95;	Conservative 0;	Mismatches 100;	Indels 0;	Gaps 0;

Qy	441	aaatttattttc	taaccaccac	catgac	tgcata	atttc	caaac	ctgac	agcag	agtc	cat	500	
Db	1161	ATAATCAAGT	TTTTGGAA	CAACTG	TGAC	ACGCA	TCATC	CAACCA	ATTC	TG	GAC	AGTTATC 1107	
Qy	501	gcgaggtg <th>gaagc</th> <th>gaagac</th> <th>tgaat</th> <th>caat</th> <th>caatcc</th> <th>ccagtt</th> <th>gtcg</th> <th>agtc</th> <th>gaag</th> <th>agatg 560</th>	gaagc	gaagac	tgaat	caat	caatcc	ccagtt	gtcg	agtc	gaag	agatg 560	
Db	1101	AACATTCG <td>CCAATC<td>TCCTCT</td><td>TAT</td><td>TAAAC<td>CCCCCT<td>CTTGAT<td>TTCT</td><td>AGT</td><td>TAAT</td><td>AC</td><td>GAGAG 104</td></td></td></td></td>	CCAATC <td>TCCTCT</td> <td>TAT</td> <td>TAAAC<td>CCCCCT<td>CTTGAT<td>TTCT</td><td>AGT</td><td>TAAT</td><td>AC</td><td>GAGAG 104</td></td></td></td>	TCCTCT	TAT	TAAAC <td>CCCCCT<td>CTTGAT<td>TTCT</td><td>AGT</td><td>TAAT</td><td>AC</td><td>GAGAG 104</td></td></td>	CCCCCT <td>CTTGAT<td>TTCT</td><td>AGT</td><td>TAAT</td><td>AC</td><td>GAGAG 104</td></td>	CTTGAT <td>TTCT</td> <td>AGT</td> <td>TAAT</td> <td>AC</td> <td>GAGAG 104</td>	TTCT	AGT	TAAT	AC	GAGAG 104
Qy	561	acactg <th>aaagt</th> <th>atgc</th> <th>gtat</th> <th>cat</th> <th>atg</th> <th>atc</th> <th>atc</th> <th>atc</th> <th>atc</th> <th>atc</th> <th>atc 620</th>	aaagt	atgc	gtat	cat	atg	atc	atc	atc	atc	atc	atc 620
Db	1041	TATTTCA <td>CAACAT</td> <td>CAGAT</td> <td>GATTTT</td> <td>ACTT</td> <td>CTC</td> <td>ATTAAT</td> <td>CAAT</td> <td>CAGAT</td> <td>CTTT</td> <td>CAC</td> <td>TGAGAC 982</td>	CAACAT	CAGAT	GATTTT	ACTT	CTC	ATTAAT	CAAT	CAGAT	CTTT	CAC	TGAGAC 982
Qy	621	ctacaa <th>ttgt</th> <th>tttt</th> <th>635</th> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	ttgt	tttt	635								
Db	981	TAAT	TAATTT	GTCT	967								

## RESULT 12

```
US-09-053-375B-600/c
; Sequence 600, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
```



```
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053,375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 600
; LENGTH: 10495
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-053-375B-600
```

```
Query Match 3.4%; Score 34.4; DB 5; Length 10495;
Best Local Similarity 53.8%; Pred. No. 71;
Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
```

```
Qy 587 tcattacatacgtacacatacacaatgtaaccctacaattgttttttgagcagagt 646
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2476 TGAGTTGCACTGCTGAAGAGACATGCTGAACCTAAGACTTTTCTTTCCATTGGAAA 2417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 647 ggtgtgtcttttttttttttttttttttttttttttttttttttttttttttt 706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2416 CTTTGGACTTTTGTGTCATTTTTCACATGTTCTTCTGTGGCAGACATTAACAGAAAGTCAG 2357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 707 atgatggtcgg 718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2356 CTTCTCGGGCTG 2345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

## RESULT 13

```
PCT-US02-09105-482
; Sequence 482, Application PC/TUS0209105
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS951PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09105
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 779
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 482
; LENGTH: 15932
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US02-09105-482
```

```
Query Match 3.4%; Score 34.4; DB 1; Length 15932;
Best Local Similarity 57.4%; Pred. No. 84;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```
Qy 546 gtgcgaagaagatgacatgaagatgcgtatcagattcattacatacatgtaca 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3900 gtgtataaaaaaagaacaaatgtatcatgtataatgattccaatttatacata 3959
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 606 aatacataatgtacacctacaattgttttttgagcagagtggtgtg 653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3960 cacacataaagagactgttaagaacttgggagcgagcgagtggtg 4007
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 14
PCT-US02-09370-1049
; Sequence 1049, Application PC/TUS0209370
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
```

```
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS954PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09370
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 1834
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1049
; LENGTH: 15932
; TYPE: DNA
; ORGANISM: Homo sapiens
; PCT-US02-09370-1049
```

```
Query Match 3.4%; Score 34.4; DB 1; Length 15932;
Best Local Similarity 57.4%; Pred. No. 84;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```
Qy 546 gtgcgaagaagatgacatgaagatgcgtatcagattcattacatacatgtaca 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3900 gtgtataaaaaaagaacaaatgtatcatgtataatgattccaatttatacata 3959
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 606 aatacataatgtacacctacaattgttttttgagcagagtggtgtg 653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3960 cacacataaagagactgttaagaacttgggagcgagcgagtggtg 4007
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

## RESULT 15

```
US-10-105-299-6415
; Sequence 6415, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; PRIOR APPLICATION removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6415
; LENGTH: 15932
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-105-299-6415
```

```
Query Match 3.4%; Score 34.4; DB 7; Length 15932;
Best Local Similarity 57.4%; Pred. No. 84;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```
Qy 546 gtgcgaagaagatgacatgaagatgcgtatcagattcattacatacatgtaca 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3900 gtgtataaaaaaagaacaaatgtatcatgtataatgattccaatttatacata 3959
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 606 aatacataatgtacacctacaattgttttttgagcagagtggtgtg 653
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3960 cacacataaagagactgttaagaacttgggagcgagcgagtggtg 4007
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Search completed: September 12, 2002, 09:12:57
Job time: 11174 sec
```



Gencore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 10:08:12 ; Search time 5258.2 Seconds  
(without alignments)  
4118.621 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_4500\_5500  
Sequence: 1 tactagcttaagattccca.....tgtattaaccagttgaca 1001

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 1081749327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_NA\_Main:\*

1:	/cgn2_6/ptodata/1/pna/US06.COMB.seq:*
2:	/cgn2_6/ptodata/1/pna/US06.COMB.seq:*
3:	/cgn2_6/ptodata/1/pna/US07.COMB.seq:*
4:	/cgn2_6/ptodata/1/pna/US08.COMB.seq:*
5:	/cgn2_6/ptodata/1/pna/US081.COMB.seq:*
6:	/cgn2_6/ptodata/1/pna/US082.COMB.seq:*
7:	/cgn2_6/ptodata/1/pna/US083.COMB.seq:*
8:	/cgn2_6/ptodata/1/pna/US084.COMB.seq:*
9:	/cgn2_6/ptodata/1/pna/US085.COMB.seq:*
10:	/cgn2_6/ptodata/1/pna/US086.COMB.seq:*
11:	/cgn2_6/ptodata/1/pna/US087.COMB.seq:*
12:	/cgn2_6/ptodata/1/pna/US088.COMB.seq:*
13:	/cgn2_6/ptodata/1/pna/US089.COMB.seq:*
14:	/cgn2_6/ptodata/1/pna/US090.COMB.seq:*
15:	/cgn2_6/ptodata/1/pna/US091.COMB.seq:*
16:	/cgn2_6/ptodata/1/pna/US092.COMB.seq:*
17:	/cgn2_6/ptodata/1/pna/US093.COMB.seq:*
18:	/cgn2_6/ptodata/1/pna/US094.COMB.seq:*
19:	/cgn2_6/ptodata/1/pna/US095A.COMB.seq:*
20:	/cgn2_6/ptodata/1/pna/US095B.COMB.seq:*
21:	/cgn2_6/ptodata/1/pna/US095C.COMB.seq:*
22:	/cgn2_6/ptodata/1/pna/US095D.COMB.seq:*
23:	/cgn2_6/ptodata/1/pna/US096A.COMB.seq:*
24:	/cgn2_6/ptodata/1/pna/US096B.COMB.seq:*
25:	/cgn2_6/ptodata/1/pna/US096C.COMB.seq:*
26:	/cgn2_6/ptodata/1/pna/US096D.COMB.seq:*
27:	/cgn2_6/ptodata/1/pna/US096E.COMB.seq:*
28:	/cgn2_6/ptodata/1/pna/US097A.COMB.seq:*
29:	/cgn2_6/ptodata/1/pna/US097B.COMB.seq:*
30:	/cgn2_6/ptodata/1/pna/US097C.COMB.seq:*
31:	/cgn2_6/ptodata/1/pna/US098A.COMB.seq:*
32:	/cgn2_6/ptodata/1/pna/US098B.COMB.seq:*
33:	/cgn2_6/ptodata/1/pna/US098C.COMB.seq:*
34:	/cgn2_6/ptodata/1/pna/US099A.COMB.seq:*
35:	/cgn2_6/ptodata/1/pna/US099B.COMB.seq:*
36:	/cgn2_6/ptodata/1/pna/US099C.COMB.seq:*
37:	/cgn2_6/ptodata/1/pna/US100.COMB.seq:*
38:	/cgn2_6/ptodata/1/pna/US101.COMB.seq:*
39:	/cgn2_6/ptodata/1/pna/US6000.COMB.seq:*
40:	/cgn2_6/ptodata/1/pna/US6001.COMB.seq:*
41:	/cgn2_6/ptodata/1/pna/US6002.COMB.seq:*
42:	/cgn2_6/ptodata/1/pna/US6003.COMB.seq:*
43:	/cgn2_6/ptodata/1/pna/US6004.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	11463	19 US-09-508-377-10	Sequence 10, Appl
2	141.2	14.1	104364	20 US-09-534-859-337	Sequence 337, App
3	141.2	14.1	104364	31 US-09-803-736-337	Sequence 337, App
4	133	13.3	416	21 US-09-606-977-62778	Sequence 62778, A
5	131.6	13.1	102897	31 US-09-803-736-1401	Sequence 1401, Ap
6	128.2	12.8	2853	11 US-08-726-705-3	Sequence 3, Appl
7	128.2	12.8	2853	23 US-09-609-040-3	Sequence 42943, A
8	127.8	12.8	2368	24 US-09-620-392-42943	Sequence 20657, A
9	127.8	12.8	2368	28 US-09-702-134-20657	Sequence 91233, A
10	127.8	12.8	2368	31 US-09-815-264-91233	Sequence 26244, A
11	126.6	12.6	400	22 US-09-572-409-26244	Sequence 16419, A
12	123.6	12.3	345	31 US-09-804-730-16419	Sequence 14485, A
13	123.6	12.3	345	17 US-60-189-657-14485	Sequence 4612, Ap
14	111.6	11.1	23449	17 US-09-330-889A-2	Sequence 385160, Sequence 385160,
15	109	10.9	2655	33 US-09-887-272A-4612	Sequence 54797, A
16	109	10.9	3017	25 US-09-654-617-385160	Sequence 98960, A
17	109	10.9	3017	27 US-09-684-016-385160	Sequence 1, Appl
18	105.8	10.6	724	33 US-09-865-439A-54797	Sequence 13950, A
19	105.8	10.6	724	33 US-60-207-458-98860	Sequence 75452, A
20	105.8	10.6	2446	29 US-09-731-168-9	Sequence 28820, A
21	105.4	10.5	2559	30 US-09-792-127-1	Sequence 1027, Ap
22	105.4	10.5	3039	30 US-09-792-127-1	Sequence 1243, Ap
23	103.2	10.3	21414	24 US-09-620-392-13950	Sequence 454409, Sequence 454409,
24	103.2	10.3	21414	28 US-09-702-134-28820	Sequence 454409, Sequence 454409,
25	103.2	10.3	21414	31 US-09-815-264-77542	Sequence 1612, Ap
26	102.4	10.2	1642	71 US-60-325-448-1027	
27	101.8	10.2	378	32 US-09-850-147-1243	
28	101.8	10.2	378	32 US-60-202-213-1240	
29	101.8	10.2	384	25 US-09-654-617-454409	
30	101.8	10.2	384	27 US-09-684-016-454409	
31	101.4	10.1	388	27 US-09-692-257A-1612	

32	101.4	10.1	388	55	US-60-162-747-1455	Sequence 1455, A
33	100.8	10.1	686	53	US-09-865-4398-553497	Sequence 53497, A
34	100.8	10.1	686	59	US-60-207-4558-97660	Sequence 97660, A
35	100.8	10.1	2087	14	US-09-091-052-9	Sequence 9
36	100.8	10.1	2087	16	US-09-257-894-9	Sequence 9, Appl1
37	100.8	10.1	2165	14	US-09-091-052-8	Sequence 9, Appl1
38	100.8	10.1	2165	16	US-09-357-894-8	Sequence 8, Appl1
39	100.8	10.1	2324	25	US-09-654-617-259798	Sequence 8, Appl1
40	100.8	10.1	2324	27	US-09-684-016-259798	Sequence 259798, A
41	100.8	10.1	2640	18	US-09-402-052-55	Sequence 259798, A
42	100.8	10.1	2665	14	US-09-091-052-1	Sequence 55, Appl1
43	100.8	10.1	2665	16	US-09-257-894-1	Sequence 1, Appl1
44	100.8	10.1	2725	19	US-09-731-894-1	Sequence 11, Appl1
45	100.8	10.1	2725	24	US-09-625-406-14	Sequence 14, Appl1

## ALIGNMENTS

```

RESULT 1
US-09-508-377-10
Sequence 10, Application US/09508377
GENERAL INFORMATION:
APPLICANT: KALEEN, ZHONGYILI
APPLICANT: MORELL, MATTHEW
APPLICANT: RAHMAN, SADEOUR
TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS
FILE REFERENCE: 094270/0126
CURRENT APPLICATION NUMBER: US/09/508,377
CURRENT FILING DATE: 2000-06-09
PRIORITY APPLICATION NUMBER: AU PP 2509
PRIORITY FILING DATE: 1998-03-20
PRIORITY APPLICATION NUMBER: PCT/AU98/00743
PRIORITY FILING DATE: 1998-09-11
PRIORITY APPLICATION NUMBER: AU PP 9108
PRIORITY FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 11463
TYPE: DNA
ORGANISM: Triticum tauschii
US-09-508-377-10

```

Query Match	100.0%;	Score 1001;	DB 19;	Length 11463;
Best Local Similarity	100.0%;	Pred. No. 7.8e-231;		
Matches 1001; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	1	taaagcttaagattccccccttagatgtaagaanaatttgtattggagcgcctccagca	60
Db	4500	taccagcttaagattccccccttagatgtaagaanaatttgtattggagcgcctccagca	4555
QY	61	agccattctaccattatataagagagagacaaggggggggggggggggttccc	120
Db	4560	agccattctaccattatataagagagagacaaggggggggggggggggttccc	4619
QY	121	ttcatattctcgcgagcgattccaanaactccatctgtctcgggtgaactcgcagg	180
Db	4620	ttcatattctcgcgagcgattccaanaactccatctgtctcgggtgaactcgcagg	4679
QY	181	atctcccatatgaagagagatatgtaattctctgttaacctacttggaaacttgatct	240
Db	4680	atctcccatatgaagagagatatgtaattctctgttaacctacttggaaacttgatct	4739
QY	241	tgaagatcgcctatatatatctatcatcacaaacttaagagatgcatttgaatttaag	300
Db	4740	tgaagatcgcctatatatatctatcatcacaaacttaagagatgcatttgaatttaag	4799
QY	301	tgtgattctgcacaaggaaccgaagataaatttatatgcctaatttagagatgagttgtg	360
Db	4800	tgtgattctgcacaaggaaccgaagataaatttatatgcctaatttagagatgagttgtg	4859

QY	361	ccaagaatlaaaagcttggatacaaatgagcagacataaaggcaatccaaggaagctta	420
Db	4860	ccaagaatlaaaagcttggatacaaaagcagcagatgaataaggcaatccaaggaagctta	4913
QY	421	tactatgcaagcttgggtatctcaacaatccatctttttctgtatacactcttaacca	480
Db	4920	tactatgcaagcttgggtatctcaacaatccatctttttctgtatacactcttaacca	4972
QY	481	tttggagctatataccccaatgcttcacatcaataaataattggatataatcccttat	540
Db	4980	tttggagctatataccccaatgcttcacatcaataaataattggatataatcccttat	5033
QY	541	tagatatatgataaactacaccttagatctctgaaaagaatcttttatgtgtgtgct	600
Db	5040	tagatatatgataaactacaccttagatctctgaaaagaatcttttatgtgtgtgct	5099

Oy	661	gacttaaatccttgatccgatagacacagacgttggttggttcttatgtatatt	720
Db	5160	gacttaaaatccttgatccgatagacacagacgttggttggttcttatgtatatt	5215
Oy	721	gtccatagtaattagtcacatattgaattagctgtttacgttttatcttgatctcaa	780
Db	5220	gtccatagtaattagtcacatattgaattagctgtttacgttttatcttgatctcaa	5275
Oy	781	agggaattcagcgcaattatgatcattgtgcacaaagctaaagatgycgaagtgaattgt	840
Db	5280	agggaattcagcgcaattatgatcattgtgcacaaagctaaagatgycgaagtgaattgt	5335
Oy	841	caaatctcaggtgycgtataggaanaattgycacaaactgagcagggcaaaaataaattt	900
Db	5340	caaatctcaggtgycgtataggaanaattgycacaaactgagcagggcaaaaataaattt	5395
Oy	901	cccatctcaaaatgycagagcccatcgcgcgaataattttccattcctataaattgtgcta	960
Db	5400	cccatctcaaaatgycagagcccatcgcgcgaataattttccattcctataaattgtgcta	5455
Oy	961	cgtgactcttttttctcagatgatataaacacagcttgaca	1001
Db	5460	cgtgactcttttttctcagatgatataaacacagcttgaca	5500

RESULT 2  
US-09-534-859-337  
; Sequence 337, Application US/09534859

1 APPLICANT: Bush, David F.  
 2 APPLICANT: Last, Robert L.  
 3 APPLICANT: Levin, Irena M.  
 4 APPLICANT: Norris, Susan R.  
 5 APPLICANT: Parnell, Laurence D.  
 6 APPLICANT: Rounsley, Steven D.  
 7 APPLICANT: Wiegand, Roger C.  
 8  
 9 TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF  
 10  
 11 FILE REFERENCE: 38-10(15493)B  
 12  
 13 CURRENT APPLICATION NUMBER: US/09/534,859  
 14  
 15 CURRENT FILING DATE: 2000-03-29  
 16  
 17 NUMBER OF SEQ ID NOS: 1127  
 18  
 19 SEQ ID NO 337  
 20  
 21 LENGTH: 104364  
 22  
 23 TYPE: DNA  
 24  
 25 ORGANISM: Arabidopsis thaliana  
 26  
 27 US-09-534-859-337

Query Match	14.1%;	Score 141.2;	DB 20;	length 104364;
Best Local Similarity	62.2%;	Pred. No. 9.3e-23;		
Matches 270;	Conservative 0;	Mismatches 138;	Indels 26;	Gaps 2

QY 313 caggaaccgaagataatcattatgtctaattttagggatgaggtgttgcagaagaattaaa 372

```
Db 16557 caggaaacaaagataatacatatgccaacttagagatgagtacttccocgataaa 16616
Qy 373 aggccttgatacatgcatgacataatgagcaatccagagacatctacataatgaagc 432
Db 16617 aagctgagctataatgctgtagataatgagccatcaagagcatgctcctatagcca 16676
Qy 433 ttggagatcaacaacatccattttctctgtatcaccttccaccattggagat 492
Db 16677 ttggagat-----cttcccttccagataataaatgtctactactatcatt 16726
Qy 493 acatccaaatgcttcacatgacataaataatgagataataccttattagataatgt 552
Db 16727 aatgaatacctgttaaaagcatttggattctctcagagtc-----gt 16770
Qy 553 acaactaaccttagtatcttgcgaaagaatcatttatgttgtgtgtccaggtac 612
Db 16771 aaacactattgtctccttgatgaatgagcatgtgacttgcataatgcaatgaggtat 16830
Qy 613 catgtactaattttttgcacaaagtagcgttttggaaacctcagagactaaatcc 672
Db 16831 cagtgacaaatcttctgcacaaagtagcgcgtttggaacacctgataccttaact 16890
Qy 673 ttgacgatagagacatgagctgtgtgtgtcttctatagatattgttcaatgaa 732
Db 16891 ttgatagaacaaagcatgagctaggtctgtgtctctgataatgttgacaaggtat 16950
Qy 733 ttagtccaatttaa 746
Db 16951 tggcttctctttaa 16964
```

## RESULT 3

```
US-09-803-736-337
; Sequence 337, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 337
; LENGTH: 104364
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-337
```

```
Query Match 14.1%; Score 141.2; DB 31; Length 104364;
Best Local Similarity 62.2%; Pred. No. 9.3e-23;
Matches 270; Conservative 0; Mismatches 138; Indels 26; Gaps 2;
```

```
Qy 313 caggaaacgaaagataatcatatgcttaatttagagtagaggtgtgtgcaagaattaa 372
Db 16557 caggaaacaaagataatacatatgccaacttagagagatgagtacttccocgataaa 16616
Qy 373 aggccttgatacatgcatgacataatgagcaatccagagacatctacataatgaagc 432
Db 16617 aagctgagctataatgctgtagataatgagccatcaagagcatgctcctatagcca 16676
Qy 433 ttggagatcaacaacatccattttctctgtatcaccttccaccattggagat 492
Db 16677 ttggagat-----cttcccttccagataataaatgtctactactatcatt 16726
```

```
Qy 493 acatccaaatgcttcacatgacataaataatgagataataccttattagataatgt 552
Db 16727 aatgaatacctgtttaaagcatttggattctctcagagtc-----gt 16770
Qy 553 acaactaaccttagtatcttgcgaaagaatcatttatgttgtgtgtccaggtac 612
Db 16771 aaacactattgtctccttgatgaatgagcatgtgacttgcataatgcaatgaggtat 16830
Qy 613 catgtactaattttttgcacaaagtagcgttttggaaacctcagagactaaatcc 672
Db 16831 cagtgacaaatcttctgcacaaagtagcgcgtttggaacacctgataccttaact 16890
Qy 673 ttgacgatagagacatgagctgtgtgtgtcttctatagatattgttcaatgaa 732
Db 16891 ttgatagaacaaagcatgagctaggtctgtgtctctgataatgttgacaaggtat 16950
Qy 733 ttagtccaatttaa 746
Db 16951 tggcttctctttaa 16964
```

## RESULT 4

```
US-09-606-977-62778
; Sequence 62778, Application US/09606977
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15877)B
; CURRENT APPLICATION NUMBER: US/09/606,977
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 82359
; SEQ ID NO 62778
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: unsure at all n locations
US-09-606-977-62778
```

```
Query Match 13.3%; Score 133; DB 23; Length 416;
Best Local Similarity 66.3%; Pred. No. 1.4e-21;
Matches 240; Conservative 0; Mismatches 105; Indels 17; Gaps 3;
```

```
Qy 387 tgcagtgacgataatgagcaatccagagacatctacataatgcaagcttgagttacaa 446
Db 1 tgcagtgacgataatgagcaatctacataatgcaagcttgagttacaa 60
Qy 447 ca-----atcatttttctgtatcactc-----ttcacccattggagct 490
Db 61 cttaaatgataccatccatttcttggttctctgttggtttcaatccatccgaactc 120
Qy 491 ttacatcctaattgtcttcacatgacata-aatatlttgataataatccttattagatat 549
Db 121 ttgcttctcctaattgttgcataaagagacatctgataatctgctcattatttctgtag 180
Qy 550 agtaacaactaaccttgatctcgcgaaagaatcatttatgttgtgtgtgtccag 609
Db 181 ccataaacaataatagtntcttactatagatagntctgtatatttccatnccag 240
Qy 610 tacatgtactaattttttgcacaaagtagcgttttggaaacctcagagactaa 669
Db 241 taccatgttaagaattttttgcccagaagtagcgttttggagctccagagactaa 300
Qy 670 tcttgatgataagacatgagctgtgtgtgtctgttcttataatgatatgttcaatg 729
Db 301 tctctattgataaagcagatgagctgtgtgtgtctgtatgatatgttcatang 360
Qy 730 ta 731
Db 361 ta 362
```









```

QY 316 gaaccgaagataatcatatgcttaatttagagatgaggtgttggccaagaattaaag 375
      |||||||
Db 45 gaaccgaagataatcatatgcttaatttagagatgaggtgttggccaagaattaaag 104
QY 376 ctgtgatacatgcaatgcaagataatgcaatcagagacatcactacataagcctt 435
      |||||||
Db 105 ctgtgatacatgcaatgcaagataatgcaatcagagacatcactacataagcctt 164
QY 436 gggatcacaacatcatttttc 461
      |||||
Db 165 gggatcacaatcatttttttc 190

```

## RESULT 14

```

US-09-330-889A-2
; Sequence 2, Application US/09330889A
; GENERAL INFORMATION:
; APPLICANT: Kim, Kyung-Nam
; APPLICANT: Gullitnan, Mark A.
; TITLE OF INVENTION: Expression Control Elements from Genes
; FILE REFERENCE: Penn State 1465
; CURRENT APPLICATION NUMBER: US/09/330,889A
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/089,049
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: US 60/089,050
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 23449
; TYPE: DNA
; ORGANISM: Zea mays
US-09-330-889A-2

```

```

Query Match 11.1%; Score 111.6; DB 17; Length 23449;
Best Local Similarity 81.6%; Pred. No. 7.9e-16;
Matches 129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

```

```

QY 309 tgacacgaacgaagataatcgaatcttaatttagagatgaggtgttggccaagaat 368
      |||||||
Db 7585 tgattggaacgaagataatcgaatcttaatttagagatgaggtgttggccaagaat 7644
QY 369 taaaggttgatgatacatgcaatgcaagataatgcaatcagagacatcactacatg 428
      |||||
Db 7645 aaaaacttgatgatacatgcaatgcaagataatgcaatcagagacatcactacatg 7704
QY 429 aagcttggtatcacaatcattttttctgtat 466
      |||||||
Db 7705 aagcttggtatcacaatcattttttctgtat 7742

```

## RESULT 15

```

US-09-887-272A-4612
; Sequence 4612, Application US/09887272A
; GENERAL INFORMATION:
; APPLICANT: Hou, Yu-Ming
; APPLICANT: Quan, Sheng
; APPLICANT: Chang, Hur-Song
; APPLICANT: Zhu, Tong
; APPLICANT: Whitlam, Steve
; APPLICANT: Goff, Steve
; APPLICANT: Glazebrook, Jane
; APPLICANT: Chen, Wenguiang
; APPLICANT: Katagiri, Fumiki
; APPLICANT: Xie, Zhiyi
; APPLICANT: Tao, Yi
; APPLICANT: Zou, Guangzhou
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST

```

```

; TITLE OF INVENTION: PATHOGENS
; FILE REFERENCE: 1360.003US2
; CURRENT APPLICATION NUMBER: US/09/887,272A
; PRIOR FILING DATE: 2001-06-23
; CURRENT APPLICATION NUMBER: 60/213,634
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/214,926
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/261,320
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/264,353
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/273,879
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 09/887,271
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 6813
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4612
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-272A-4612

```

```

Query Match 10.9%; Score 109; DB 33; Length 2655;
Best Local Similarity 83.2%; Pred. No. 1.6e-15;
Matches 124; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```

```

QY 313 caagaacgaagataatcatalcttaatttagagatgaggtgttggccaagaattaa 372
      |||||||
Db 1093 ccgaacgaagataatcatalcttaatttagagatgaggtgttggccaagaattaa 1152
QY 373 aagcttggtatcacaatcattttttctgtat 461
      |||||||
Db 1153 aagcttggtatcacaatcattttttctgtat 1212
QY 433 ttgggtatcacaatcattttttctgtat 461
      |||||||
Db 1213 ttgggtatcacaatcattttttctgtat 1241

```

```

Search completed: September 12, 2002, 10:11:16
Job time: 9624 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 10:11:16 : Search time 5258.2 Seconds  
(without alignments)  
4118.621 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_9500\_10500  
Perfect score: 1001  
Sequence: 1 tccgacatagacacagcatga.....ctggagcgaatagctttttg 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US080\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq:\*  
9: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq:\*  
10: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq:\*  
11: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq:\*  
12: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq:\*  
13: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq:\*  
14: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq:\*  
15: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq:\*  
16: /cgn2\_6/ptodata/1/pna/US092\_COMB.seq:\*  
17: /cgn2\_6/ptodata/1/pna/US093\_COMB.seq:\*  
18: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:\*  
19: /cgn2\_6/ptodata/1/pna/US095\_COMB.seq:\*  
20: /cgn2\_6/ptodata/1/pna/US096\_COMB.seq:\*  
21: /cgn2\_6/ptodata/1/pna/US097\_COMB.seq:\*  
22: /cgn2\_6/ptodata/1/pna/US098\_COMB.seq:\*  
23: /cgn2\_6/ptodata/1/pna/US099\_COMB.seq:\*  
24: /cgn2\_6/ptodata/1/pna/US100\_COMB.seq:\*  
25: /cgn2\_6/ptodata/1/pna/US101\_COMB.seq:\*  
26: /cgn2\_6/ptodata/1/pna/US102\_COMB.seq:\*  
27: /cgn2\_6/ptodata/1/pna/US103\_COMB.seq:\*  
28: /cgn2\_6/ptodata/1/pna/US104\_COMB.seq:\*  
29: /cgn2\_6/ptodata/1/pna/US105\_COMB.seq:\*  
30: /cgn2\_6/ptodata/1/pna/US106\_COMB.seq:\*  
31: /cgn2\_6/ptodata/1/pna/US107\_COMB.seq:\*  
32: /cgn2\_6/ptodata/1/pna/US108\_COMB.seq:\*  
33: /cgn2\_6/ptodata/1/pna/US109\_COMB.seq:\*  
34: /cgn2\_6/ptodata/1/pna/US110\_COMB.seq:\*  
35: /cgn2\_6/ptodata/1/pna/US111\_COMB.seq:\*  
36: /cgn2\_6/ptodata/1/pna/US112\_COMB.seq:\*  
37: /cgn2\_6/ptodata/1/pna/US113\_COMB.seq:\*  
38: /cgn2\_6/ptodata/1/pna/US114\_COMB.seq:\*  
39: /cgn2\_6/ptodata/1/pna/US115\_COMB.seq:\*  
40: /cgn2\_6/ptodata/1/pna/US116\_COMB.seq:\*  
41: /cgn2\_6/ptodata/1/pna/US117\_COMB.seq:\*  
42: /cgn2\_6/ptodata/1/pna/US118\_COMB.seq:\*  
43: /cgn2\_6/ptodata/1/pna/US119\_COMB.seq:\*

44: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq:\*  
45: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq:\*  
46: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq:\*  
47: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq:\*  
48: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq:\*  
49: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq:\*  
50: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq:\*  
51: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq:\*  
52: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq:\*  
53: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq:\*  
54: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq:\*  
55: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq:\*  
56: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq:\*  
57: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq:\*  
58: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq:\*  
59: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq:\*  
60: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq:\*  
61: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq:\*  
62: /cgn2\_6/ptodata/1/pna/US6023\_COMB.seq:\*  
63: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq:\*  
64: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq:\*  
65: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq:\*  
66: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq:\*  
67: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq:\*  
68: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq:\*  
69: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq:\*  
70: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq:\*  
71: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq:\*  
72: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq:\*  
73: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq:\*  
74: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq:\*  
75: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	11463	19 US-09-508-377-10	Sequence 10, Appl
2	223.2	22.3	23449	17 US-09-330-889A-2	Sequence 13, Appl
3	187.2	18.7	21414	24 US-09-620-352-13950	Sequence 28820, A
4	187.2	18.7	21414	31 US-09-702-134-28820	Sequence 77542, A
5	187.2	18.7	21414	28 US-09-815-264-77542	Sequence 3410, Ap
6	136.4	13.6	1119	71 US-60-325-448-3410	Sequence 3, Appl
7	134.8	13.5	2853	11 US-08-726-705-3	Sequence 4612, Ap
8	128.8	12.9	2655	33 US-09-887-212A-4612	Sequence 91564, A
9	122	12.2	420	17 US-09-304-517A-91564	Sequence 21714, A
10	122	12.2	421	36 US-09-985-678-91564	Sequence 42955, A
11	122	12.2	421	17 US-09-394-745-21714	Sequence 91564, A
12	122	12.2	421	22 US-09-565-306-42955	Sequence 2236, Ap
13	122	12.2	499	17 US-09-371-146A-91564	Sequence 1807, Ap
14	122	12.2	531	32 US-09-845-526A-2236	Sequence 204100, Ap
15	122	12.2	531	29 US-09-731-166-9	Sequence 204100, Ap
16	122	12.2	531	29 US-09-731-166-9	Sequence 204100, Ap
17	122	12.2	531	29 US-09-731-166-9	Sequence 204100, Ap
18	120.4	12.0	399	24 US-09-637-086A-1807	Sequence 21048, A
19	120.4	12.0	399	24 US-09-637-086A-1807	Sequence 33904, A
20	120.4	12.0	399	25 US-09-654-617-204100	Sequence 93092, A
21	120.4	12.0	399	25 US-09-654-617-204100	Sequence 385160, A
22	120.4	12.0	441	17 US-09-304-517A-93092	Sequence 485167, A
23	120.4	12.0	441	17 US-09-394-745-21048	Sequence 69979, A
24	120.4	12.0	441	22 US-09-565-306-39904	Sequence 2386, Ap
25	120.4	12.0	441	36 US-09-985-678-93092	
26	120.4	12.0	527	17 US-09-371-146A-93092	
27	118.8	11.9	3017	25 US-09-654-617-385160	
28	118.8	11.9	3017	27 US-09-684-016-385160	
29	117.8	11.8	413	17 US-09-394-745-48567	
30	117.8	11.8	413	22 US-09-565-306-69979	
31	117.8	11.8	567	32 US-09-845-526A-2386	

32	117.8	11.8	567	59	US-60-202-214-2386	Sequence 2386, Ap
33	117.2	11.7	366	32	US-09-850-147-7191	Sequence 7191, Ap
34	117.2	11.7	366	59	US-60-202-213-7178	Sequence 7178, Ap
35	117.2	11.7	368	32	US-09-850-147-4345	Sequence 4345, Ap
36	117.2	11.7	368	59	US-60-202-213-4342	Sequence 4342, Ap
37	117.2	11.7	725	25	US-09-654-617-452844	Sequence 452844, Ap
38	117.2	11.7	725	27	US-09-684-016-452844	Sequence 452844, Ap
39	115.8	11.6	444	26	US-09-667-188A-2250	Sequence 2250, Ap
40	115.8	11.6	444	54	US-60-155-006-3134	Sequence 3134, Ap
41	115.8	11.6	457	26	US-09-667-188A-1619	Sequence 1619, Ap
42	115.8	11.6	457	54	US-60-155-006-2503	Sequence 2503, Ap
43	115.6	11.5	365	32	US-09-850-147-6476	Sequence 6476, Ap
44	115.6	11.5	365	59	US-60-202-213-6463	Sequence 6463, Ap
45	110.8	11.1	291	17	US-09-304-517A-49291	Sequence 49291, A

## ALIGNMENTS

RESULT 1  
US-09-508-377-10  
; Sequence 10, Application US/09508377  
; GENERAL INFORMATION:  
; APPLICANT: KALEEN, ZHONGYILI  
; APPLICANT: MORELL, MATHEW  
; APPLICANT: RAHMAN, SADEOUR  
; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION IN PLANTS  
; FILE REFERENCE: 054270/0126  
; CURRENT APPLICATION NUMBER: US/09/508, 377  
; CURRENT FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: AU PP 2509  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: PCT/AU98/00743  
; PRIOR FILING DATE: 1998-09-11  
; PRIOR APPLICATION NUMBER: AU PP 9108  
; PRIOR FILING DATE: 1997-09-12  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 11463  
; TYPE: DNA  
; ORGANISM: Triticum tauschii  
US-09-508-377-10

Query Match 100.0%; Score 1001; DB 19; Length 11463;  
Best Local Similarity 100.0%; Pred. No. 3.2e-231;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	tcgcacatagacagcatgaacatatcaagctctctctgtgcaggatatagtatgtca	60
DB	9500	tcgcacatagacagcatgaacatatcaagctctctctgtgcaggatatagtatgtca	9559
QY	61	tggtcttgatagagcttaactcttcgcatgctgtgtgcatacataaataatgat	120
DB	9560	tggtcttgatagagcttaactcttcgcatgctgtgtgcatacataaataatgat	9619
QY	121	caggtcttcacatgggttaggtgtggaaggtcattcaattcattgggaagaatt	180
DB	9620	caggtcttcacatgggttaggtgtggaaggtcattcaattcattgggaagaatt	9679
QY	181	tgggcatctgtgcatgctttatcaacattatgcatcttcgcatgattgtgttactta	240
DB	9680	tgggcatctgtgcatgctttatcaacattatgcatcttcgcatgattgtgttactta	9739
QY	241	attgaacacatgctttcttcaacatgtatgtatgtataatcgtgtgtccaaagag	300
DB	9740	attgaacacatgctttcttcaacatgtatgtatgtataatcgtgtgtccaaagag	9799
QY	301	gaagtaactcttattacttgacagaaatgatatgatttccaaagagcccaaacct	360
DB	9800	gaagtaactcttattacttgacagaaatgatatgatttccaaagagcccaaacct	9859

QY	361	tccaaccgcaaaagtctcccttggaataacaatagttatgataaagccgcgtagat	420
DB	9860	tccaaccgcaaaagtctcccttggaataacaatagttatgataaagccgcgtagat	9919
QY	421	ttgattcttgtaagtttaagctgtgtctattacattccctactaataattttatggcaat	480
DB	9920	ttgattcttgtaagtttaagctgtgtctattacattccctactaataattttatggcaat	9979
QY	481	tattcttgtaagtaacataatgtttgttggaagaagtaacattgctttgtatgtttg	540
DB	9980	tattcttgtaagtaacataatgtttgttggaagaagtaacattgctttgtatgtttg	10039
QY	541	tagacgttaacaataagatagtgtgtgaagttgtgtatcatataaataatcatgtat	600
DB	10040	tagacgttaacaataagatagtgtgtgaagttgtgtatcatataaataatcatgtat	10099
QY	601	gcaggagatgcgaatttcttctaagatatacgtgtatgcgaagttcgatagcgatgca	660
DB	10100	gcaggagatgcgaatttcttctaagatatacgtgtatgcgaagttcgatagcgatgca	10159
QY	661	gcattcttgaggaataataggggtatgtcaactgtgttcttcttgatatacaagtcac	720
DB	10160	gcattcttgaggaataataggggtatgtcaactgtgttcttcttgatatacaagtcac	10219
QY	721	agtttaacgtcagctctcttcaagtgtgtaaaaaagtgtagaataatcctgtatgaga	780
DB	10220	agtttaacgtcagctctcttcaagtgtgtaaaaaagtgtagaataatcctgtatgaga	10279
QY	781	tgaaacctgtcgaagggagctggaattgtcttccacaaacttttcttaagttgc	840
DB	10280	tgaaacctgtcgaagggagctggaattgtcttccacaaacttttcttaagttgc	10339
QY	841	tttgtatgtatatacatatacagcactgcgaatgttaactgcagttatgacatctgaca	900
DB	10340	tttgtatgtatatacatatacagcactgcgaatgttaactgcagttatgacatctgaca	10399
QY	901	ccagatatttcacggaacaatgaggaagaatgtagatcatcctcctaaagaagattt	960
DB	10400	ccagatatttcacggaacaatgaggaagaatgtagatcatcctcctaaagaagattt	10459
QY	961	ggtattgttttcaactccactgagagaatgctttttg	1001
DB	10460	ggtattgttttcaactccactgagagaatgctttttg	10500

RESULT 2  
US-09-330-889A-2  
; Sequence 2, Application US/09330889A  
; GENERAL INFORMATION:  
; APPLICANT: Guiltinan, Mark A.  
; APPLICANT: Kim, Kyung-Nam  
; TITLE OF INVENTION: Expression Control Elements from Genes  
; TITLE OF INVENTION: Encoding Starch Branching Enzymes  
; FILE REFERENCE: Penn State 1465  
; CURRENT APPLICATION NUMBER: US/09/330, 889A  
; CURRENT FILING DATE: 1999-06-11  
; PRIOR APPLICATION NUMBER: US 60/089, 049  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: US 60/089, 050  
; PRIOR FILING DATE: 1998-06-12  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 23449  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-330-889A-2

Query Match 22.3%; Score 223.2; DB 17; Length 23449;  
Best Local Similarity 58.6%; Pred. No. 6.3e-43;  
Matches 649; Conservative 0; Mismatches 333; Indels 126; Gaps 10;

QY	19	gaacatcacgaagctctcttttggcagagatgataagatttaagttcaatgctctgtgaag-cct	77
D	17812	gaacctcaactgtgtctatttattatgacagagatgataagatttaagttcaatgctctgtgaagctt	17871
QY	78	caacctctgcattatctgtgcataagacattacataaattgatgcaggtctgcacacatg	137
D	17872	caacctcaactcattatctgtgtgataagacattacataaattgatgattgaattacataaagg	17931
QY	138	gtttatgtgtgaaagcctacctaacttcattcatggaagaagatttggcagctctgttcagt	197
D	17932	gtttatgtgtgaaagcctacctaacttcattcatggaagaagatttggcagctctgttcagta	17991
QY	198	ctttacaacattatgtcattctgcattgatgtgatttactgtla-----ttgaac	248
D	17992	tttaactactattgtttcttaacttcacttcgtttgagctttatagaacagttacattccaac	18051
QY	249	catgtttcttctcaactgtatgtatattatgtaactgtttgtgtccaaaggagaaagttaa	308
D	18052	aattactctgcaattacattcctttgtttagttatattatgattgtgtgaagacttgagttact	18111
QY	309	cttccatttactctgagaaatggaatgataatttccaaaggcccccaaaactcttccaaag	368
D	18112	cttctattatttttgagaaatgataagatttccaaagggtcccgaaagacttccaaagt	18171
QY	369	gcaaaagtctccctcgtgaaataacataatgtatgatataatgcgcgcgtatgattcctt	428
D	18172	gt-aagttattctcagggaaataacacaagttatgacaatgtcgtcgaagatttgactgt	18230
QY	429	gtaagtttagtgtgcatactacattccctcaactagatccttatggcca-----	478
D	18231	gtaaacttctttgtattgtgcaaaagttccaaagttgtatttacttlaacttlaaccattgcca	18290
QY	479	-----ttattctctgataaatacattgaatttg	507
D	18291	tgctttaatcagcaaggtgcacattataatagtctctctcttcttaattatattagcagttcc	18350
QY	508	ttaggaagaatcaacattgtcttgt--agttttagagcgttlaacataagatagtgtg	565
D	18351	agaagtaaaaataattactacactctgttlaaagtttcttataatatagtcgttgttg	18410
QY	566	agagttgttgcatttaaaaaatacatgattttt-----gcaggagatagcagattt	618
D	18411	tcaaatctgcattatcaggaattgtgttatttlaactgttcaggggtgtagcagacta	18470
QY	619	tctttagatactgtgtatgcaagagttcogatcaggaatgcagcactctgaaagaaaata	678
D	18471	tctttagatactgtgtatgcaagagtttgcacaggaatgcacactcttggcaaaaata	18530
QY	679	tggggtatg-----tcaacgtgttgccttgttgcataaacaagtacagttt	725
D	18531	tgaagtaagttccttttttacttttttattgtgttcgtcgaaggtttccacaacattcat	18590
QY	726	aacgtcagttcttctaagttgtaaaaaaagttagaattatctctg-----taagt	777
D	18591	attgttgtgcattctactgttgaagttcaatttlaaaaaaacattctcctcagtttactg	18650
QY	778	agattgaanaactgtgcagaaggcggagctgtgaattgt-----tttcacaaaactatttc	832
D	18651	agcttttaagcaatgaaggtttcatatagaattcttcatgattgttgcatacaactcttag	18710
QY	833	ttaagttcgttgtatgtata-----	853
D	18711	gtatttcatatgattcatatagaattactctggaagacagccataatggtlaacgaaaaat	18770
QY	854	catataccaagcctigacaaatgtaactgcgaatttaagaattctgagccacagatatttca	913
D	18771	ttctgtatgaaatttgcgtgtgtaattgtgaattcattgaacatcgtgataccagatatttcc	18830
QY	914	cggaaacatgaagaagataaagttgatactccccaagaagagatttggatttttcc	973
D	18831	cggaaacatgaagaagataaagttgtgttcggaaaaggagatttggattttgtgttc	18890
QY	974	aacttccacttggagcaaatagcttttttg	1001

```

||||| | |||||
Db 18891 aaclccactgcacacaagcattcttg 18918

RESULT      3
US-09-620-392-13950/c
; Sequence 13950, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
SEQ ID NO 13950
; LENGTH: 21414
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-13950

Query Match          18.7%; Score 187.2; DB 24; Length 21414;
Best Local Similarity 71.5%; Pred. No. 3.1e-34;
Matches 306; Conservative 0; Mismatches 108; Indels 14; Gaps 4

Oy 19 gaacatcacagctctctcttggcgaggataatgatattcaatggtctcgtaagctc- 77
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10778 GAACCTCACATCTGTACTTATGCAAGGATATGATATTTTATGGCTGCAGACGCCG 10719

Oy 78 caacccttcgcatgatcgtagcatagcataataaatgatcaagctgttcacatag 137
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10718 CAACACCTTAGCATTTGATCGTGGAATTAAGCATTCATTAATGATTAATCACAATGG 10659

Oy 138 gtttagtgtgaagcctactctaactcatatggaatatgattgggcalctcgtgcagt 197
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10658 GGTTAGGAGGAGAAGCGTATCTTAACCTTATGSGAAATGAGTCCGACATCCGTGGAGA 10599

Oy 138 ctccaacattatgcatctcgcgatgatgtgattactgt-----aattgaac 248
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10598 TTTAATAACAACAACACACACAGCTTTGTTCTTATTTTATACACTTCATTTATGTA 10539

Oy 249 catgctttcttcacatgttatgtatataatcgtgtgtccaaagagaagttaa 308
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10538 CATGCTCTTGTCACTAGTATTTTCTGTGGAAACAAATTGGCT---AAAGGCATTTGAT 10482

Oy 309 cttctattacttggcagaatgatagatttccaagagagccacaactcctccaacg 368
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10481 CTACCTTTTATTTTGCAGATAGATTGATTTTCCAAAGACCTCCACAAAGTCCAATG 10422

Oy 369 gcaaaagtctcccccggaataatacatagttatgataaatgcgcgtatgattgacct 428
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10421 GTAAA-TTTCATCCGAGGGAATTAACAACAAGTATGATTAATAATGCCGTGGAAGATTGACCTG 10363

Oy 429 gtaagttc 436
    ||||| | | |
Db 10362 GTAACCTT 10355

RESULT      4
US-09-702-134-28820/c
; Sequence 28820, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
```

```

:
: CURRENT APPLICATION NUMBER: US/09/702,134
: CURRENT FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 52202
: SEQ ID NO 28820
: LENGTH: 21414
: TYPE: DNA
: ORGANISM: Oryza sativa
: US-09-702-134-28820

```

Query Match	18.7%;	Score 187.2;	DB 28;	Length 21414;
Best Local Similarity	71.5%;	Pred. No. 3.1e-34;		
Matches 306;	Conservative	0;	Mismatches 108;	Indels 14;
				Gaps 4;

QY 19 gaacatatcaagctcctcttgtgcaagatagtatgatccatgcctctgataggctt-77  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 10778 GAACCTCACATACTTAATTATGACAGATATGTATGATTTTATGCGCTCGACAGACGG 10719

QY 78 caactcttcgcatcttgatcgtgycatagcaattacataaaatgatacagcttgcaccatgg 137  
|||||  
Db 10718 CAACACCTAGACATTGATCGTGAATAGACATTGCATAAAAAGATTAGACTATACCAATGG 10659

QY 138 gtttaagtggtgaagcgtatctttaactlcatggaagtlltggcgatccctgctcgt 197  
+  
Db 10658 GGTTAGGAGAGAAGCCTATCTTAACCTTATGGGAATGAGTTCCGACATCCTGGTGAGA 10599

Qy 198 cttacaacattatctgcatcctgcgatgttgatttaccyt-----aatltgaac 248  
|| || || | | | | | | | |  
Db 10538 TTATATAACAATCACACACCAGTTGTTCCTATTATATCAACTTCTCATTTGATGA 10539

QY 249 catgctttcttcacatgttatgtatatgtaatctgttgcctccaagaggaagttaa 308  
|| || || || || || || || || || || || || || || || || || || || ||  
Db 10538 CAATGCTTCTGTACACTAGTATTTCGTGGAAACAATTGGCTT--AAAGCCATTGAT 10482

```
QY      309 cttcatttacttggcagaatgatatagattlccaagagcccaacaactcttccaaccy   368
        || | |||| | ||||||| ||||| ||||| ||||| ||||| |
Db    10481 CTACCTTTTATTTTCGAGAATGATGTGATTTCCAAAGAGCTCCACAAGTACTCTCCAATG 10422
```

Qy 369 gcaaaagtcccccctggaataacaatagttatgataaaagcgcggtagattgatctt 428  
+ ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 10421 GTAA-TTCATCCCAAGGAAATACACAAGTTAATGAATAAACCGTCAAGATTGCCTG 10363

|    |       |          |       |
|----|-------|----------|-------|
| QY | 429   | gtaagttt | 436   |
|    |       |          |       |
| Db | 10362 | GTAACCTT | 10355 |

RESULT 5  
US-09-815-264-77542/C  
Sequence 77542 Annotation NC\_00915364

```

; GENERAL INFORMATION:
;
; APPLICANT: Boukharov, Andrey A.
;
; APPLICANT: Cao, Yongwei
;
; APPLICANT: Dotson Stanton B

```

; APPLICANT: Koshi, Jeffrey M.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Liu, Jingtong  
 ; APPLICANT: McElroy, James

: APPLICANT: Wu, Wei  
 :  
 : TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
 :  
 : FILE REFERENCE: 38-21(51237)G  
 :  
 : CURRENT APPLICATION NUMBER: US/09/815 264

: CURRENT FILING DATE: 2001-03-23  
 : PRIOR APPLICATION NUMBER: US 09/620,392  
 : PRIOR FILING DATE: 2000-07-19  
 : PRIOR APPLICATION NUMBER: US 09/702 134

```

:      PRIOR FILING DATE: 2000-10-31
:
:      NUMBER OF SEQ ID NOS: 109669
:
:      SEQ ID NO 77542
:
:      LENGTH: 21414

```

```

; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-77542

```

|                       |                 |                    |            |               |
|-----------------------|-----------------|--------------------|------------|---------------|
| Query Match           | 18.7%;          | Score 187.2;       | DB 31;     | Length 21414; |
| Best Local Similarity | -71.5%;         | Pred. No. 3.1e-34; |            |               |
| Matches 306;          | Conservative 0; | Mismatches 108;    | Indels 14; | Gaps 4;       |

10778 GAACTCACATCTACTGTAATTATCCAGATATGTAATGATTTATGCTCTGCACAGACCGG 10719

78 caacttcgcattgcatcgtgcataagcattacaataatgatccggctgtcaacatyg 137  
y ||||| | ||||||||| ||||||| ||||||| ||||| ||||| |||||  
10718 CAACACTAGCATTTGATCGTGGATAAGCATTGCCTAAATGATTTAGACTTCATCACCAATGG 10659

b  
10658 GGTTAGGAGGAAGCCTATCTTAACTTTATGGAAATAGATTCCGACATCCTGTGAGA 10599

b  
10598 TTTAATACAAATCCACAACCAGTGTGTTCTATTATCAACTTCATCTTGATGA 10539

b  
249 catgccttcttcacacattgtatgatatglaactcgttgctcccaaggaagaattaa 308  
|| || | ||| | ||| | ||| | ||| |  
10538 CAATGCTTCTTCACACTAGTATTTCTGGAAACAATTGGCTT---AAAAGCCATTGTAT 10482

[illegible]

y 369 gcaaaatctccccctgtaataataacaatagttatgataaatgcgcgtgatattgactc 428  
| | | | | | | | | | | | | | | | | | | |  
b 10421 GTAAATTTCAATCCACGGGAATTACAACAGTATGATAAATGCCCGGAAGATTTGCACCTG 10363

Y 429 gtaagltt 436  
|||||  
b 10362 GTACTTT 10355

RESULT 6  
S-60-325-448-3410  
COUNTS 3410 ANALYST: JEN: JN/603354/48

GENERAL INFORMATION:  
APPLICANT: Budworth, P.R.  
APPLICANT: Moughamer, T.G.  
FIRM OR INSTITUTION: Consulting Engineers

FILE REFERENCE: 1360.026PRV  
CURRENT APPLICATION NUMBER: US/60/325,448  
CURRENT FILING DATE: 2001-09-26  
NUMBER OF SEQ. ID NOS.: 4709

SEQ ID NO 3410  
LENGTH: 1119  
TYPE: DNA  
ORFANTRY: Escherichia coli

S-60-325-448-3410

Querry: Watch 13 69; 60000 136 A; DB 71; Length 1110.

Best Local Similarity - 98.7%; Pred. No. 2.5e-22;  
Matches 148; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

b  
|||  
aggaatgtatgatcatctgactgtatagaacctcaaccccgcatgcattggca 288  
109 taccagttaaatataaacacgcccttatcacacaattttagtttgttaaaggcatttccta 161

289 tagcattacataaataatgcacagcttgcaccatggtttaggtcgttgaaagcctatccta 348  
163 acttataatgggaaatgagctttatgggcatcctat 101

349 actcatggaatgatttggcatcctg 378



```
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/214,926
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/261,320
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/264,353
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/273,879
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 09/887,271
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 6813
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4612
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-272A-4612
```

Query Match 12.2%; Score 128.8; DB 33; Length 2655;  
Best Local Similarity 84.8%; Pred. No. 2.2e-20;  
Matches 156; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

```
QY 9 agaacagcatgaacatacaagctctcttgcgagatattgatttcacatgctctg 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1975 acaacagcagcatatctcaactctcttcattcatgagatattgatttcacatg 2034

QY 69 gatagc-cttcaactctctgcattgctgcgacatagacataaataagcctt 127
    || || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2035 gacagaccttcaacacctgcgattgacgctgagatagacataaataagcctt 2094

QY 128 gtccacctgggtttagtggtgtaagctatcttaactcattgaggaattggtgcat 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2095 gtccacctgggtttagtggtgtaagctatcttaactcattgaggaattggtgcat 2154

QY 188 cctg 191
    |||||
DB 2155 cctg 2158
```

```
RESULT 10
US-09-304-517A-91564
; Sequence 91564, Application US/09304517A
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)B
; CURRENT APPLICATION NUMBER: US/09/304,517A
; CURRENT FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 91564
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Zea mays
US-09-304-517A-91564
```

Query Match 12.2%; Score 122; DB 17; Length 420;  
Best Local Similarity 92.7%; Pred. No. 5.7e-19;  
Matches 139; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
QY 43 aggatatgatatttcacgctctgataag-cttcaactcttcgcatgtatcgagca 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 aggatatgatatttcacgctctgataag-cttcaactcttcgcatgtatcgagca 139

QY 102 tagcattacataaataatgacagctgttcaccatggttttagtggtgtaaggtcatctta 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 tagcattacataaataatgacagctgttcaccatggttttagtggtgtaaggtcatctta 199

QY 162 acttcattggaatgagtttggcctcctg 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

DB 200 atttcattggaatgagtttggcctcctg 229

```
RESULT 11
US-09-985-678-91564
; Sequence 91564, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 91564
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Zea mays
US-09-985-678-91564
```

Query Match 12.2%; Score 122; DB 36; Length 420;  
Best Local Similarity 92.7%; Pred. No. 5.7e-19;  
Matches 139; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
QY 43 aggatatgatatttcacgctctgataag-cttcaactcttcgcatgtatcgagca 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 aggatatgatatttcacgctctgataag-cttcaactcttcgcatgtatcgagca 139

QY 102 tagcattacataaataatgacagctgttcaccatggttttagtggtgtaaggtcatctta 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 tagcattacataaataatgacagctgttcaccatggttttagtggtgtaaggtcatctta 199

QY 162 acttcattggaatgagtttggcctcctg 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 atttcattggaatgagtttggcctcctg 229
```

```
RESULT 12
US-09-394-745-21714
; Sequence 21714, Application US/09394745
; GENERAL INFORMATION:
; APPLICANT: Fisher, Dane K.
; APPLICANT: Lajoudi, Raghuath V.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WI
; FILE REFERENCE: 38-21(15454)B
; CURRENT APPLICATION NUMBER: US/09/394,745
; CURRENT FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 57264
; SEQ ID NO 21714
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3062-002-Q1-K2-E10
US-09-394-745-21714
```

Query Match 12.2%; Score 122; DB 17; Length 421;  
Best Local Similarity 92.7%; Pred. No. 5.7e-19;  
Matches 139; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
QY 43 aggatatgatatttcacgctctgataag-cttcaactcttcgcatgtatcgagca 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 81 aggatatgatatttcacgctctgataag-cttcaactcttcgcatgtatcgagca 140

QY 102 tagcattacataaataatgacagctgttcaccatggttttagtggtgtaaggtcatctta 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 141 tagcattacataaataatgacagctgttcaccatggttttagtggtgtaaggtcatctta 200

QY 162 acttcattggaatgagtttggcctcctg 191
```







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 09:13:22 : Search time 643.61 seconds  
(without alignments)  
4921.411 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_9500\_10500  
Perfect score: 1001  
Sequence: 1 tcgcagacatagaacagcatga.....ctggagcaatagctttttt 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2084544 seqs, 1582152599 residues  
Total number of hits satisfying chosen parameters: 4169088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_New:\*  
1: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                 | Description        |
|------------|-------|-------------|--------|-----------------------|--------------------|
| 1          | 130   | 13.0        | 2554   | 7 US-10-162-948-1     | Sequence 1, Appl   |
| 2          | 122   | 12.2        | 3144   | 8 US-60-391-781-685   | Sequence 685, App  |
| 3          | 110.8 | 11.1        | 2443   | 8 US-60-368-387-1     | Sequence 1, Appl   |
| 4          | 110.8 | 11.1        | 2443   | 8 US-60-381-534-1     | Sequence 1, Appl   |
| 5          | 110.8 | 11.1        | 2919   | 8 US-60-391-781-709   | Sequence 709, App  |
| 6          | 102.6 | 10.2        | 786    | 8 US-60-391-781-2614  | Sequence 2614, App |
| 7          | 93.6  | 9.4         | 3003   | 7 US-10-056-454A-12   | Sequence 12, Appl  |
| 8          | 92.8  | 9.3         | 2529   | 7 US-10-056-454A-17   | Sequence 17, Appl  |
| 9          | 92    | 9.2         | 3033   | 7 US-10-056-454A-14   | Sequence 14, Appl  |
| 10         | 90.4  | 9.0         | 2576   | 7 US-10-056-454A-16   | Sequence 16, Appl  |
| 11         | 90.4  | 9.0         | 2578   | 7 US-10-056-454A-19   | Sequence 19, Appl  |
| 12         | 90.4  | 9.0         | 2975   | 7 US-10-056-454A-13   | Sequence 13, Appl  |
| 13         | 90.4  | 9.0         | 3231   | 7 US-10-056-454A-18   | Sequence 18, Appl  |
| 14         | 90    | 9.0         | 1867   | 6 US-10-219-999-3997  | Sequence 3997, App |
| 15         | 67.4  | 6.7         | 2096   | 6 US-10-213-073-149   | Sequence 149, App  |
| 16         | 67.4  | 6.7         | 2913   | 1 PCT-US02-18947-442  | Sequence 442, App  |
| 17         | 67.4  | 6.7         | 2913   | 1 PCT-US02-18947-442  | Sequence 442, App  |
| 18         | 67.4  | 6.7         | 2955   | 1 PCT-US02-25766-8907 | Sequence 8907, App |
| 19         | 60.8  | 6.1         | 4563   | 1 US-10-110-777-1     | Sequence 1, Appl   |
| 20         | 59.2  | 5.9         | 2324   | 6 US-10-219-999-16451 | Sequence 16451, A  |
| 21         | 59.2  | 5.9         | 2551   | 8 US-60-368-387-2     | Sequence 2, Appl   |
| 22         | 59.2  | 5.9         | 2551   | 8 US-60-381-534-2     | Sequence 2, Appl   |
| 23         | 54.8  | 5.5         | 2007   | 7 US-10-179-131-2516  | Sequence 2516, App |
| 24         | 53.6  | 5.4         | 1941   | 8 US-60-360-039-33970 | Sequence 33970, A  |
| 25         | 53.4  | 5.3         | 2638   | 8 US-60-360-039-27656 | Sequence 27656, A  |

|    |      |     |         |                       |                    |
|----|------|-----|---------|-----------------------|--------------------|
| 26 | 51.2 | 5.1 | 2115    | 8 US-60-360-039-25407 | Sequence 25407, A  |
| 27 | 50.8 | 5.1 | 189     | 5 US-09-975-254-16738 | Sequence 16738, A  |
| 28 | 50.6 | 5.1 | 184     | 5 US-09-975-254-29078 | Sequence 29078, A  |
| 29 | 42.4 | 4.2 | 5525    | 7 US-10-172-086-8     | Sequence 8, Appl   |
| 30 | 41.6 | 4.2 | 2025    | 5 US-09-540-2098-375  | Sequence 375, App  |
| 31 | 41.6 | 4.2 | 96589   | 6 US-10-035-832-1226  | Sequence 1226, App |
| 32 | 41.6 | 4.2 | 96589   | 6 US-10-052-482-214   | Sequence 214, App  |
| 33 | 40.8 | 4.1 | 6246    | 7 US-10-172-086-28    | Sequence 28, Appl  |
| 34 | 40.6 | 4.1 | 2132    | 7 US-10-099-278-39    | Sequence 39, Appl  |
| 35 | 39.4 | 3.9 | 2140405 | 7 US-10-027-632-76212 | Sequence 76212, A  |
| 36 | 38.8 | 3.9 | 1669    | 5 US-09-591-466C-1    | Sequence 1, Appl   |
| 37 | 38.6 | 3.9 | 500     | 5 US-09-991-936-601   | Sequence 601, App  |
| 38 | 38.6 | 3.9 | 839     | 7 US-10-155-881-14448 | Sequence 14448, A  |
| 39 | 38.4 | 3.8 | 2601    | 7 US-10-179-131-2345  | Sequence 2345, App |
| 40 | 37.8 | 3.8 | 2635    | 7 US-10-105-299-8524  | Sequence 8524, App |
| 41 | 37.6 | 3.8 | 4238    | 6 US-10-205-219-152   | Sequence 152, App  |
| 42 | 37.4 | 3.7 | 310     | 7 US-10-102-524-1616  | Sequence 1616, App |
| 43 | 37.2 | 3.7 | 1191139 | 5 US-09-811-352B-1    | Sequence 1, Appl   |
| 44 | 37.2 | 3.7 | 1191139 | 7 US-10-140-924-1     | Sequence 1, Appl   |
| 45 | 37.2 | 3.7 | 1691139 | 7 US-10-067-514-1     | Sequence 1, Appl   |

## ALIGNMENTS

```

RESULT 1
US-10-162-948-1
; Sequence 1, Application US/10162948
; GENERAL INFORMATION:
; APPLICANT: Jane, Jay-Jin
; APPLICANT: Mou, Beiquan
; APPLICANT: Jansson, Christer
; APPLICANT: Sun, Chuanxin
; TITLE OF INVENTION: TRANSGENIC CORN PLANTS HAVING SEEDS WITH MODIFIED
; TITLE OF INVENTION: CORNSARCH CHARACTERISTICS AND METHOD OF MAKING THE
; FILE REFERENCE: 480927.90081
; CURRENT APPLICATION NUMBER: US/10/162,948
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/295,649
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2554
; TYPE: DNA
; ORGANISM: Hordeum vulgare
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(2208)
US-10-162-948-1

Query Match 13.0%; Score 130; DB 7; Length 2554;
Best Local Similarity 96.0%; Pred. No. 5.5e-23;
Matches 144; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 43 aggatatgatattcatgctctgtagatag-gcttaactcttgcatgcatgagca 101
|||||
DB 1568 aggatatgatattcatgctctgtagatagatgacccctcgcattgcatgagca 1627
|||||

QY 102 taacattacataaagatcagcttgtaacattgattagtgtaggaagctactta 161
|||||
DB 1628 taacattacataaagatcagcttgtaacattgattagtgtaggaagctactta 1687
|||||

QY 162 actcatggaatgagtttgagcatcctg 191
|||||
DB 1688 attcatggaatgagtttgagcatcctg 1717
|||||

RESULT 2
US-60-391-781-685
; Sequence 685, Application US/60391781

```

```

: GENERAL INFORMATION:
: APPLICANT: Laurie, Cathy C.
: TITLE OF INVENTION: Nucleic Acid Molecules Associated With Oil in Plants
: FILE REFERENCE: 38-77(52900)B
: CURRENT APPLICATION NUMBER: US/60/391,781
: CURRENT FILING DATE: 2002-06-25
: PRIOR APPLICATION NUMBER: 60/365,301
: PRIOR FILING DATE: 2002-03-15
: NUMBER OF SEQ ID NOS: 3034
: SEQ ID NO 685
: LENGTH: 3144
: TYPE: DNA
: ORGANISM: Zea mays
: US-60-391-781-685

```

|                       |                |                   |          |             |
|-----------------------|----------------|-------------------|----------|-------------|
| Query Match           | 12.2%          | Score 123         | DB 8     | length 3144 |
| Best Local Similarity | 92.7%          | Pred. No. 6.3e-21 |          |             |
| Matches 139           | Conservative 0 | Mismatches 10     | Indels 1 | Gaps 1      |

Qy 43 aggaatatgatgattcattcattgctctggaatg-cttcaactcttcgaattgacgttgca 101  
|||||  
Db 2107 aggaatatgatgattcattcattgctctggaacggccttcaacgctcgcgatcgatcgttggaa 2166  
|||||

|    |  |      |   |      |
|----|--|------|---|------|
| OY |  | 102  | tacgatttcataaaatgatcaggcttgcaccatgggttttagtggtgaagcatctta     | 161  |
|    |  |      |   |      |
| Db |  | 2167 | tacgattcataaaaatgatcaggcttgcacaatgggttttaggagcgcaagcgtatatcaa | 2226 |

| QY | 162  | acttcattggaatgagtttggcattcctg | 191  |
|----|------|-------------------------------|------|
| Db | 2227 | atttcattggaatgagtttggcattcctg | 2256 |

```

RESULT      3
US-60-368-387-1
: Sequence 1, Application US/60368387
: GENERAL INFORMATION:
: APPLICANT: Pearlstein, Richard W.
: APPLICANT: Broglio, Karen E.
: APPLICANT: Hines, Christopher F.
: TITLE OF INVENTION: Maltize Starch Containing Elevated Amounts of Actual Amylose
: FILE REFERENCE: BB1510PRV
: CURRENT APPLICATION NUMBER: US/60/368,387
: CURRENT FILING DATE: 2002-03-27
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO: 1
: LENGTH: 2443
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(2443)
: OTHER INFORMATION: SBEII modified region
: US-60-368-387-1

```

|                           |        |            |              |    |    |        |       |
|---------------------------|--------|------------|--------------|----|----|--------|-------|
| Query Match               | 11.1%  | Score      | 110.8;       | DB | 8; | Length | 2443; |
| Best Local Similarity     | 88.0%; | Pred.      | No. 4.1e-18; |    |    |        |       |
| Matches 132; Conservative | 0;     | Mismatches | 17;          |    |    | Indels | 1;    |
|                           |        |            |              |    |    | Gaps   | 1;    |

QY 43 aggatatgatgatcctcctgataag-gcttcaacctcttcgcatgctgagca 101  
|||||  
Db 1411 aggatattgatcctcctgataagccttcaacctctacattgactgagca 1470  
|||||

Oy 102 tagcatcataaataatgacagcctgtcaacatggtttaagtgtgtaagcctatccta 161  
 |||||  
 Db 1471 tagcatcataaataatgacagcctgtcaacatggtttaagtgtgtaagcctatccta 1530

QY 162 acttcattggaatgagtttggcacctg 191  
| ||||| ||||| |||||  
Db 1531 atttcataagaaataagtttaacacatccta 1560

```

RESULT      4
US-60-381-534-1
; Sequence 1, Application US/60381534

```

```

1 GENERAL INFORMATION:
2 APPLICANT: Pearlstein, Richard W.
3 APPLICANT: Brogile, Karen E.
4 APPLICANT: Hines, Christopher F.
5 TITLE OF INVENTION: Maize Starch Containing Elevated Amounts of Actual Amylose
6 FILE REFERENCE: BB1510PRV1
7 CURRENT APPLICATION NUMBER: GS/50/381,534
8 CURRENT FILING DATE: 2002-05-16
9 NUMBER OF SEQ ID NOS: 6
10 SOFTWARE: PatentIn version 3.1

```

```

: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)..(2443)
: OTHER INFORMATION: SBFI modified region
US-60-381-534-1

```

|                           |       |                   |                   |
|---------------------------|-------|-------------------|-------------------|
| Query Match               | 11.1% | DB 8:             | length 2443:      |
| Best Local Similarity     | 88.0% | Pred. NC.4.1e-18: |                   |
| Matches 132; Conservative | 0;    | Mismatches 17;    | Indels 1; Gaps 1; |

Qy 43 aggatatgatgatattcaatgctctgatag -gcttcaactcttcgatattgcgtggca 101  
|||||  
Db 1411 agaatatgatgatattcaatgcccctgatagaccttcaactctacattgatcgtggga 1470  
|||||

Qy 102 tagcattacataaaatgcatcgctgtcacacatgggtttagtggtgaagccatctta 161  
|||||  
Db 1471 taccattacataaatgatgattagacttatcaccaatggtgttgaggagaagggctatctta 1530  
|||||

| QY | 162  | acttcattggaatgagtttggcaccctg | 191  |
|----|------|------------------------------|------|
|    |      |                              |      |
| Db | 1531 | atttcattggaatgagtttggcaccctg | 1560 |

```

RESULT      5
US-60-391-781-709
: Sequence 709, Application US/60391781
: GENERAL INFORMATION:
: APPLICANT: Lartie, Cathy C.
: TITLE OF INVENTION: Nucleic Acid Molecules Associated With Oil in Plants
: FILE REFERENCE: 38-77(52900)B
: CURRENT APPLICATION NUMBER: US/60/391,781
: CURRENT FILING DATE: 2002-06-25
: PRIOR APPLICATION NUMBER: 60/365,301
: PRIOR FILING DATE: 2002-03-15
: NUMBER OF SEQ ID NOS: 3034
: SEQ ID NO 709
: LENGTH: 2919
: TYPE: DNA
: ORGANISM: Zea mays
: US-60-391-781-709

```

|                           |        |                    |           |              |
|---------------------------|--------|--------------------|-----------|--------------|
| Query Match               | 11.1%  | Score 110.8;       | DB 8;     | Length 2919; |
| Best Local Similarity     | 88.0%; | Pred. No. 4.3e-18; |           |              |
| Matches 132; Conservative | 0;     | Mismatches 17;     | Indels 1; | Gaps 1;      |

43 aggaatatgatgattcaagctctgataag-gcttcaacctctcgattgactgtgca 101  
 1906 aggaatatgatgattcaagctctgataag-gcttcaacctctcgattgactgtgca 1965

QY 102 tagcattacataaaatgacagcctgtccacatgggtttagtggtagagctatctta 161  
|||||  
Db 1966 tacacattccataaatgaattgaacattatcaccaaatattgaacgaagaagcattccta 2025  
|||||

| Qy | 162  | acttcattggaatgagtttggcattcctg | 191  |
|----|------|-------------------------------|------|
|    |      |                               |      |
| Db | 2026 | atttcattggaatgagtttggcattcctg | 2055 |

```

RESULT      6
US-60-391-781-2614/C
; Sequence 2614, Application US/60391781
; GENERAL INFORMATION:
; APPLICANT: Laurie, Cathy C.
; TITLE OF INVENTION: Nucleic Acid Molecules Associated With Oil in Plants
; FILE REFERENCE: 38-77(52900)B
; CURRENT APPLICATION NUMBER: US/60/391,781
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/365,301
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 3034
; SEQ ID NO 2614
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(786)
; OTHER INFORMATION: unsure at all n locations
US-60-391-781-2614

```

|                           |       |                   |           |            |
|---------------------------|-------|-------------------|-----------|------------|
| Query Match               | 10.2% | Score 102.6       | DB 8      | Length 786 |
| Best Local Similarly      | 80.5% | Pred. No. 3.6e-16 |           |            |
| Matches 120; Conservative | 0;    | Mismatches 29;    | Indels 0; | Gaps 0;    |

[illegible]

RESULT 7  
US-10-056-454A-12  
Sequence 12, Application US/10056454A  
GENERAL INFORMATION:  
APPLICANT: National Starch and Chemical Investment Holding Corporation  
TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: National Starch and Chemical Investment Holding Corporation  
STREET: 1000 Unidigma Blvd.  
CITY: Newcastle  
STATE: Delaware  
COUNTRY: United States of America  
ZIP: 19720  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/056,454A  
FILING DATE: 25-Jun-2002  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3003 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

```

; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-056-454A-12

```

| Query Match           | 9.4%   | Score 93.6      | DB 7          | Length 3003 |
|-----------------------|--|-----------------|---------------|-------------|
| Best Local Similarity | 78.1%  | Pred. No. 1e-13 |               |             |
| Matches 125           | Conservative   | 0               | Mismatches 34 | Indels 1    |
|                       |  |                 |               | Gaps        |
| Qy                    | 33 cctcttgacagatagatgaattcaatcagcgtctgatatag-gcttcaactcttcgcaat      | 91              |               |             |
|                       |  |                 |               |             |
|                       |  |                 |               |             |
| Db                    | 2067 CTGATGGACAAGGATATGATGATTTTATGGGCTCTGGATGAGACGCTCAACATCATTAATA   | 2126            |               |             |
| Qy                    | 92 gatcgtgcgaatgacataataaatagatcagcgttgcacacatggtgttagtggtgaa        | 151             |               |             |
|                       |  |                 |               |             |
|                       |  |                 |               |             |
| Db                    | 2127 GATCGTGGGATAGCATATACCAAGAAGATTTAGGCTCTGTACTATATGGGATTGAGAGACGAA | 2186            |               |             |
| Qy                    | 152 ggcatacttaactcagggaaatggtttggcattccg                             | 191             |               |             |
|                       |  |                 |               |             |
|                       |  |                 |               |             |
| Db                    | 2187 GGGTACCTAAATTTTCATGGGAAATGAAATTCGGCCACCCG                       | 2226            |               |             |

RESULT 8  
US-10-056-454A-17  
Sequence 17, Application US/10056454A .  
GENERAL INFORMATION:  
APPLICANT: National Starch and Chemical Investment Holding Corporation  
TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: National Starch and Chemical Investment Holding Corporation  
STREET: 1000 Unidigma Blvd.  
CITY: Newcastle  
STATE: Delaware  
COUNTRY: United States of America  
ZIP: 19720  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/056,454A  
FILING DATE: 25-Jun-2002  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2529 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-10-056-454A-17

| Query Match | Similarity | 9.3%  | Score 92.8        | DB 7          | Length 2529 |
|-------------|------------|---|-------------------|---------------|-------------|
| Best Local  | Similarity | 76.9%   | Pred. No. 1.6e-13 |               |             |
| Matches     | 123        | Conservative  | 2                 | Mismatches 34 | Indels 1    |
|             |            |   |                   | Gaps          | 1           |
| QY          | 33         | ctctcttgcagagatatgtagtattcatagcctctgtagtag-gcttcaactcttcgacatt      | 91                |               |             |
|             |            |   |                   |               |             |
| Db          | 1838       | CTGAATGGACAAAGATATGATATTATTATGGCTCTGGATAGACCGCAACATCAATTAAATA       | 1897              |               |             |
| QY          | 92         | gatacgtgcacatgacatcatcataaataatgatacagcctgtgtcacaattggttaagtgtgaa   | 151               |               |             |
|             |            |   |                   |               |             |
| Db          | 1898       | GATCGTGGGATATGCACTTGGCACAAAGATGATATTAGCTTGTAACTATGAGGATTAAGAGAGAGAA | 1957              |               |             |
| QY          | 152        | ggcattcttaactcatgaggaataatgaatttggagcattccgt                        | 191               |               |             |
|             |            |   |                   |               |             |
| Db          | 1958       | GGGTACTTAATTTTCATGTGGAAATGATATTCGGCCACCCCTG                         | 1997              |               |             |

RESULT 9  
 US-10-056-454A-14  
 ; Sequence 14, Application US/10056454A

RESULT 9  
US-10-056-454A-14  
; Sequence 14, Application US/10056454A

GENERAL INFORMATION:  
APPLICANT: National Starch and Chemical Investment Holding Corporation  
TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESS: National Starch and Chemical Investment Holding Corporation  
STREET: 1000 Unigema Blvd.  
CITY: Newcastle  
STATE: Delaware  
COUNTRY: United States of America  
ZIP: 19720  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/056,454A  
FILING DATE: 25-Jun-2002  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3033 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 145..2790  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-10-056-454A-14

```

Query Match 9.2%; Score 92; DB 7; Length 3033;
Best Local Similarity 77.5%; Pred. No. 2.6e-13;
Matches 124; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 33 cctctgtgcagaaatgatgatcctgcctcgtgataag-gcctcaactctgcgcat 91
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2068 CTGATGGCAACAGATATGATGATTTTATGGCTCTGGATAGACCGTCAACATCATTA 2127

QY 92 gatcgtgcatagcattacataaataatgatacagcctgtgtacacatgagtggtgaa 151
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2128 GATCGTGGGATPGCATTCGCAACAGATGATTAGCGCTTGACTGATGGATGAGGAGGAA 2187

QY 152 ggcatacttaactcatgaggaaatgattggtgcacctcgt 191
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2188 GGGTACTCTAATTTTCATGGGAAATGATTCGGCCACCCTG 2227

RESULT 10
US-10-056-454A-16
: Sequence 16, Application US/10056454A
:
: GENERAL INFORMATION:
: APPLICANT: National Starch and Chemical Investment Holding Corporation
: TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESS: National Starch and Chemical Investment Holding Corporation
: STREET: 1000 Uniigma Blvd.
: CITY: Newcastle
: STATE: Delaware
: COUNTRY: United States of America
: ZIP: 19720
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/056,454A
: FILING DATE: 25-Jun-2002
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
:

```

```

:      LENGTH: 2576 base pairs
:      TYPE: nucleic acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      SEQUENCE DESCRIPTION: SEQ ID NO: 16
US-10-056-454A-16

```

|                       |              |                    |                |              |
|-----------------------|--------------|--------------------|----------------|--------------|
| Query Match           | 9.0%;        | Score 90.4;        | DB 7;          | Length 2576; |
| Best Local Similarity | 76.9%;       | Pred. No. 6.3e-13; |                |              |
| Matches 123;          | Conservative | 0;                 | Mismatches 36; | Indels 1;    |
|                       |              |                    |                | Gaps 1;      |

[illegible]

RESULT 11  
 US-10-056-454A-19  
 : Sequence 19, Application US/10056454A  
 : GENERAL INFORMATION:  
 : APPLICANT: National starch and Chemical Investment Holding Corporation  
 : TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition  
 : NUMBER OF SEQUENCES: 20  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: National Starch and Chemical Investment Holding Corporation  
 : STREET: 1000 Unigema Blvd.  
 : CITY: Newcastle  
 : STATE: Delaware  
 : COUNTRY: United States of America  
 : ZIP: 19720  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/10/056,454A  
 : FILING DATE: 25-Jun-2002  
 : INFORMATION FOR SEQ. ID. NO.: 19:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2578 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
 US-10-056-454A-19

|    | Query Match   | 9.0%: Score 90.4: DB 7   | Length 2578: |
|----|---|--------------------------|--------------|
|    | Best Local Similarity   | 76.9%: Pred. No. 6.3e13: |              |
|    | Matches 123: Conservative 0: Mismatches 36: Indels 1: Gaps 1:     |                          |              |
| QY | 33 ctcttgcgcagatcatgatattcattgcgtctcgatag-gcttcaactcttcgalt       | 91                       |              |
|    |   |                          |              |
| Db | 1845 CTGATGACCAAGATATGATATTTTATGGCTCTGATGACCGCAACATCATTTAATA      | 1904                     |              |
| QY | 92 gatcgtgcacatagcatatcatatnaaalgcagcctgttcacatggtttggtgta        | 151                      |              |
|    |   |                          |              |
| Db | 1905 GATCGTGGATAGCATTCACACAGATGATTTAGGCTTGTAACTATGATGGATAGGAGAGAA | 1964                     |              |
| QY | 152 ggcatacttaactcatatgggnaatgagtttgggcatccg                      | 191                      |              |
|    |   |                          |              |
| Db | 1965 GGGTACCTTAAATTTTCATGGGAATTAATTTGGCCACCCCG                    | 2004                     |              |

RESULT 12  
US-10-056-454A-13  
; Sequence 13, Application US/10056454A  
; GENERAL INFORMATION:  
; APPLICANT: National Starch and Chemical Investment Holding Corporation  
; TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSSEE: National Starch and Chemical Investment Holding Corporation  
; STREET: 1000 Unigema Blvd.  
; CITY: Newcastle  
; STATE: Delaware  
; COUNTRY: United States of America  
; ZIP: 19720  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/056,454A  
; FILING DATE: 25-Jun-2002  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2975 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-10-056-454A-13

Query Match 9.0%; Score 90.4; DB 7; Length 2975;  
Best Local Similarity 76.9%; Pred. No. 6.6e-13;  
Matches 123; Conservative 0; Mismatches 36; Indels 1; Gaps 1;  
QY 33 cctcttgacagatgatgatgctgctgagtag-gcttcaactcttcgcat 91  
DB 2066 CTGATGCACAGATGATGATGATTTTATGCTTGAGAGACCCCAACATCATTAATA 2125  
QY 92 gatcgtgcatagcataataatgatacagctgttcacacatggttagtgatga 151  
DB 2126 GATCGTGGGATGATGATGCACAGATGATTAAGCTTGTAACATGGGATTAGAGAGAGA 2185  
QY 152 ggcatacttaactcattcagtggaatgagttgggcatcctg 191  
DB 2186 GGGTACCTAAATTCATGGGAATGATTCGGCCACCCCTG 2225  
RESULT 13  
US-10-056-454A-18  
; Sequence 18, Application US/10056454A  
; GENERAL INFORMATION:  
; APPLICANT: National Starch and Chemical Investment Holding Corporation  
; TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSSEE: National Starch and Chemical Investment Holding Corporation  
; STREET: 1000 Unigema Blvd.  
; CITY: Newcastle  
; STATE: Delaware  
; COUNTRY: United States of America  
; ZIP: 19720  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/056,454A  
; FILING DATE: 25-Jun-2002  
; INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3231 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-10-056-454A-18

Query Match 9.0%; Score 90.4; DB 7; Length 3231;  
Best Local Similarity 76.9%; Pred. No. 6.8e-13;  
Matches 123; Conservative 0; Mismatches 36; Indels 1; Gaps 1;

QY 33 cctcttgacagatgatgatgctgctgagtag-gcttcaactcttcgcat 91  
DB 2151 CTGATGCACAGATGATGATGATTTTATGCTTGAGAGACCCCAACATCATTAATA 2210  
QY 92 gatcgtgcatagcataataatgatacagctgttcacacatggttagtgatga 151  
DB 2211 GATCGTGGGATGATGATGCACAGATGATTAAGCTTGTAACATGGGATTAGAGAGAGA 2270  
QY 152 ggcatacttaactcattcagtggaatgagttgggcatcctg 191  
DB 2271 GGGTACCTAAATTCATGGGAATGATTCGGCCACCCCTG 2310

RESULT 14  
US-10-219-999-3997  
; Sequence 3997, Application US/10219999  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jindong  
; APPLICANT: Stein, Joshua  
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-106527260C  
; CURRENT APPLICATION NUMBER: US/10/219,999  
; PRIOR FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: US 60/324,109  
; PRIOR FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: US 60/312,544  
; PRIOR FILING DATE: 2001-08-15  
; NUMBER OF SEQ ID NOS: 63520  
; SEQ ID NO 3997  
; LENGTH: 1867  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)..(1268)  
; OTHER INFORMATION:  
US-10-219-999-3997

Query Match 9.0%; Score 90; DB 6; Length 1867;  
Best Local Similarity 79.3%; Pred. No. 7.3e-13;  
Matches 119; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 43 aggatgatgatattcagctgctcgatag-gcttcaactcttcgcatgctgtgca 101  
DB 550 aggatgatgatattcagctgctcgatag-gcttcaactcttcgcatgctgtgca 609  
QY 102 tagcatatacaataatgatacagctgttcacacatggttagtgatga 161  
DB 610 tagcgttgacaataatgatacagctgttcacacatggttagtgatga 669  
QY 162 actcaatgggaatgagttgggcatcctg 191  
DB 670 atttatgggaatgagttgggcatcctg 699

RESULT 15  
US-10-213-073-149  
; Sequence 149, Application US/10213073  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PZ1ACIN  
; CURRENT APPLICATION NUMBER: US/10/213,073  
; CURRENT FILING DATE: 2002-08-07  
; PRIOR APPLICATION NUMBER: 09/760,483  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213



; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,064  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/233,063  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,397  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,399  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/232,401  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/241,808  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,826  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,786  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,221  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,475  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/231,243  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,065

Query Match 6.7%; Score 67.4; DB 6; Length 2096;  
Best Local Similarity 73.5%; Pred. No. 4.2e-07;  
Matches 86; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 75 cttcaactcttcgcatgtcggtgcatagcattacataaatgatacagctgttcacca 134  
DB 786 ctttaactcagctatgtcgttggaatacagcttcataaalgaltcgaactcaatcgc 845  
QY 135 tgggttaagtgtgaagctatcttaactcatggaataagtttggcatcctg 191  
DB 846 atgggctgtgtgagaagctatctcaattcatggtgaatgaatttggcatcctg 902



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 01:43:32 : Search time 3917.59 Seconds  
(without alignments)  
5341.686 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_1\_1000  
Perfect score: 1000  
Sequence: 1 agaacacccatcatttaga.....acactcccgctgggtccct 1000

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_ph:\*  
7: gb\_pl:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Query Match | Length | DB ID | Description |
|------------|-------------|--------|-------|-------------|
|------------|-------------|--------|-------|-------------|

|    |      |       |        |    |             |             |
|----|------|-------|--------|----|-------------|-------------|
| 1  | 1000 | 100.0 | 11463  | 6  | AX031271    | Sequence    |
| 2  | 1000 | 100.0 | 11475  | 8  | AF338431    | Aegilops    |
| 3  | 43.8 | 4.4   | 210476 | 10 | AC087166    | Mus muscu   |
| 4  | 43.8 | 4.4   | 225726 | 2  | AC074328    | Mus muscu   |
| 5  | 41.4 | 4.1   | 79976  | 8  | AB008264    | Arabidops   |
| 6  | 41.4 | 4.1   | 96217  | 9  | AC002410    | Human BAC   |
| 7  | 40.8 | 4.1   | 133021 | 9  | AL161624    | Human DNA   |
| 8  | 40.8 | 4.1   | 142184 | 2  | AC037428    | Human DNA   |
| 9  | 40.6 | 4.1   | 81580  | 8  | AP000736    | Homo sapi   |
| 10 | 40.6 | 4.1   | 98017  | 8  | AC027033    | Arabidops   |
| 11 | 40.6 | 4.1   | 207872 | 10 | AC074329    | Arabidops   |
| 12 | 40.0 | 4.0   | 65955  | 8  | AP003681    | Mus muscu   |
| 13 | 40.0 | 4.0   | 318221 | 2  | PFMAL13P3   | Oryza sat   |
| 14 | 39.8 | 4.0   | 7218   | 6  | I66494      | Plasmodiu   |
| 15 | 39.2 | 3.9   | 1055   | 2  | AX083745    | Sequence 14 |
| 16 | 39.2 | 3.9   | 161531 | 2  | AC079960    | Sequence    |
| 17 | 39.2 | 3.9   | 170608 | 2  | AL359372    | Homo sapi   |
| 18 | 39.2 | 3.9   | 171730 | 2  | AL589663    | Homo sapi   |
| 19 | 39.2 | 3.9   | 166527 | 9  | AC092057    | Homo sapi   |
| 20 | 39.2 | 3.9   | 170682 | 2  | AC109584    | Homo sapi   |
| 21 | 39.2 | 3.9   | 190070 | 2  | AC079032    | Homo sapi   |
| 22 | 39.2 | 3.9   | 196660 | 9  | AC007000    | Homo sapi   |
| 23 | 39.2 | 3.9   | 203426 | 2  | AL645913    | Mus muscu   |
| 24 | 38.8 | 3.9   | 140826 | 2  | OSJN00069   | Oryza sat   |
| 25 | 38.8 | 3.9   | 207166 | 2  | AC073703    | Mus muscu   |
| 26 | 38.6 | 3.9   | 102575 | 2  | AL353670    | Homo sapi   |
| 27 | 38.6 | 3.9   | 103447 | 2  | AC095853    | Homo sapi   |
| 28 | 38.6 | 3.9   | 110000 | 2  | LMFCHR36_30 | Continuatio |
| 29 | 38.6 | 3.9   | 125020 | 9  | AF429315    | Homo sapi   |
| 30 | 38.4 | 3.8   | 161031 | 9  | AC079610    | Homo sapi   |
| 31 | 38.2 | 3.8   | 43044  | 3  | AC006730    | Caenorhab   |
| 32 | 38.2 | 3.8   | 110000 | 2  | AC016150_0  | Homo sapi   |
| 33 | 38.2 | 3.8   | 44067  | 8  | AL357136    | Human DNA   |
| 34 | 38.2 | 3.8   | 57047  | 8  | LES03345    | Lycopersi   |
| 35 | 38.2 | 3.8   | 168834 | 2  | AC096964    | Rattus no   |
| 36 | 38.2 | 3.8   | 169281 | 9  | AC079232    | Mus muscu   |
| 37 | 38.2 | 3.8   | 179637 | 2  | AC007961    | Mus muscu   |
| 38 | 38.2 | 3.8   | 205884 | 2  | AC068241    | Mus muscu   |
| 39 | 37.8 | 3.8   | 33995  | 3  | CEK11E4     | Caenorhabd  |
| 40 | 37.8 | 3.8   | 34175  | 9  | CNS00YVB    | Homo sapi   |
| 41 | 37.8 | 3.8   | 111119 | 2  | AL162741    | Homo sapi   |
| 42 | 37.8 | 3.8   | 116376 | 2  | AL139287    | Homo sapi   |
| 43 | 37.8 | 3.8   | 133769 | 9  | AC010429    | Homo sapi   |
| 44 | 37.8 | 3.8   | 152116 | 9  | AC008471    | Homo sapi   |
| 45 | 37.6 | 3.8   | 1141   | 6  | AX083744    | Sequence    |

## ALIGNMENTS

| RESULT | LOCUS    | DEFINITION                         | AX031271 | SEQUENCE                           | 10 from Patent W09914314. | DNA | linear | PAT 20-SEP-2000 |
|--------|----------|------------------------------------|----------|------------------------------------|---------------------------|-----|--------|-----------------|
| 1      | AX031271 | Sequence 10 from Patent W09914314. | AX031271 | Sequence 10 from Patent W09914314. | 11463 bp                  | DNA | linear | PAT 20-SEP-2000 |
| 2      | AX031271 | Sequence 10 from Patent W09914314. | AX031271 | Sequence 10 from Patent W09914314. | 11463 bp                  | DNA | linear | PAT 20-SEP-2000 |

## KEYWORDS

## SOURCE

ORGANISM Aegilops tauschii.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Aegilops.  
1 (bases 1 to 11463)  
REFERENCE Li, Z., Morell, M. and Rahman, S.  
Regulation of gene expression in plants  
Patent: WO 9914314-A 10 25-MAR-1999;  
GOODMAN FIELDER LTD (AU) ; LI ZHONGYI (AU) ; MORELL MATTHEW (AU) ;  
RAHMAN SADEOUR (AU) ; UNIV AUSTRALIAN (AU) ; COMMW SCIENT IND RES  
ORG (AU) ; GROUPE LIMAGRAIN PACIFIC PTY L (AU)

## FEATURES

source  
1. 11463  
/organism="Aegilops tauschii"  
/db\_xref="taxon:37682"





|              |   |  |  |
|--------------|---|--|--|
| JOURNAL      | Genome Therapeutics Corporation Sequencing Center: Mouse Genome                         |  |  |
| REFERENCE    | Sequence Data   |  |  |
| AUTHORS      | Unpublished   |  |  |
| TITLE        | 2 (bases 1 to 225726)   |  |  |
| JOURNAL      | Smith,D.R.  |  |  |
|              | Direct Submission   |  |  |
|              | Submitted (23-AUG-2000) Genome Therapeutics Corporation, 100 Beaver                     |  |  |
|              | Street, Waltham, MA 02453, USA  |  |  |
|              | On Jan 15, 2002 this sequence version replaced gi:15208553.                             |  |  |
| COMMENT      | -----   |  |  |
|              | Genome Center   |  |  |
|              | Center: Genome Therapeutics Corporation   |  |  |
|              | Center code: GTC  |  |  |
|              | Web site: <a href="http://www.genomecorp.com/">http://www.genomecorp.com/</a>           |  |  |
|              | Contact: <a href="mailto:gtc-seqcenter@genomecorp.com">gtc-seqcenter@genomecorp.com</a> |  |  |
|              | ----- Project Information   |  |  |
|              | Center project name: mp042  |  |  |
|              | ----- Summary Statistics  |  |  |
|              | Sequencing vector: N/A  |  |  |
|              | Chemistry: dye-terminator Big Dye; 100% of reads  |  |  |
|              | Assembly program: Phrap; version 990315   |  |  |
|              | Consensus quality: 225213 bases at least Q40  |  |  |
|              | Consensus quality: 225480 bases at least Q30  |  |  |
|              | Consensus quality: 225564 bases at least Q20  |  |  |
|              | Insert size: 225825; sum-of-contigs   |  |  |
|              | Quality coverage: 6.6x in Q20 bases; sum-of-contigs                                     |  |  |
|              | -----   |  |  |
|              | * NOTE: This is a 'working draft' sequence. It currently                                |  |  |
|              | * consists of 1 contigs. The true order of the pieces                                   |  |  |
|              | * is not known and their order in this sequence record is                               |  |  |
|              | * arbitrary. Gaps between the contigs are represented as                                |  |  |
|              | * runs of N, but the exact sizes of the gaps are unknown.                               |  |  |
|              | * This record will be updated with the finished sequence                                |  |  |
|              | * as soon as it is available and the accession number will                              |  |  |
|              | * be preserved.   |  |  |
| FEATURES     | 1 225726: contig of 225726 bp in length.  |  |  |
| source       | Location/Qualifiers   |  |  |
|              | 1..225726   |  |  |
|              | /organism="Mus musculus"  |  |  |
|              | /strain="C57BL/6J"  |  |  |
|              | /db_xref="taxon:10090"  |  |  |
|              | /chromosome="12"  |  |  |
|              | /clone="RP23-101G12"  |  |  |
|              | /clone_lib="RPCI-23"  |  |  |
|              | 1..225726   |  |  |
| misc_feature | /note="assembly_name:Contig2  |  |  |
|              | clone_end:SP6"  |  |  |
| BASE COUNT   | 70130 a 43240 c 44470 g 67886 t   |  |  |
| ORIGIN       |   |  |  |
|              | Query Match   |  |  |
|              | Best Local Similarity 4.4%; Score 43.8; DB 2; Length 225726;                            |  |  |
|              | Matches 84; Conservative 0; Mismatches 67; Indels 0; Gaps 0;                            |  |  |
| QY           | 309 tgaacacatgatgtatcatcagccctgataagaggaacacatgcactttccct 368                           |  |  |
|              |   |  |  |
| Db           | 38555 TAAAGAAATGATGCGACAACTACCCACATCTANTGAGGACACATGAAGAAGCTTCTACG 38496                 |  |  |
| QY           | 369 ctggaatcatagtcatcactttttttaatggaagcaaggttgcacaacatgcat 428                          |  |  |
|              |   |  |  |
| Db           | 38495 AGGAAGACTCATTAGCTCAGACTGTCTACAAAAAGAAACTAGAGAAAGCACACACTAGCAG 38436               |  |  |
| QY           | 429 ttcaacaagaagaataatctccaacca 459   |  |  |
|              |   |  |  |
| Db           | 38435 CTTGACACATCTACTTAAAGCTCTGAGACAA 38405   |  |  |
| RESULT       | 5   |  |  |
| LOCUS        | AB008264 79976 bp DNA linear PLN 27-DEC-2000  |  |  |
| DEFINITION   | Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MBD2.                          |  |  |
| ACCESSION    | AB008264 BA000015   |  |  |
| VERSION      | AB008264.1 GI:2618599   |  |  |

| SOURCE                    | KEYWORDS  |
|---------------------------|---|
| ORGANISM                  | Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MBD2.   |
| REFERENCE AUTHORS         | Arabidopsis thaliana<br>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.<br>1 (sites)   |
| TITLE                     | Nakamura, Y., Sato, S., Kaneko, T., Kotani, H., Asamizu, E., Miyajima, N., and Tabata, S.   |
| JOURNAL MEDLINE REFERENCE | Structural analysis of Arabidopsis thaliana chromosome 5. III. Sequence features of the regions of 1,191,918 bp covered by seventeen physically assigned pl clones<br>DNA Res. 4 (6), 401-414 (1997)<br>98162728<br>2 (bases 1 to 79976)  |
| AUTHORS                   | Nakamura, Y.  |
| TITLE                     | Direct Submission   |
| JOURNAL                   | Submitted (24-OCT-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel:81-438-55-3935, Fax:81-438-55-3934)<br>Address for correspondence: kas@kazusa.or.jp<br>For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MBD2<br>Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.<br>The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory,<br>http://complib.oerl.gov/Grail-1.3/).<br>GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlim.zool.iastate.edu/cgi-bin/sp.cgi).<br>Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/).<br>This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions.<br>The 5' clone is MJB21 and the 3' clone is MRD20. |
| COMMENT                   | Location/Qualifiers<br>1..79976<br>/organism="Arabidopsis thaliana"<br>/strain="Columbia"<br>/db_xref="taxon:3702"<br>/chromosome="5"<br>/clone="MBD2"<br>/clone_lib="Mitsui P1"<br>complement(1..218)<br>/note="CDS is reported in Acc# AB007647<br>gene_id=MJB21.20"<br>number=1<br>/evidence=not_experimental<br>/product="v2 snrp auxiliary factor, small subunit"<br>join(5306..5761,5932..6828)<br>/note="gene_id:MBD2.2"<br>/codon_start=1<br>/evidence=not_experimental<br>/product="N-hydroxycinnamoyl/benzoyltransferase-like protein"<br>protein_id="BAB09184.1"<br>/db_xref="GI:9758571"<br>translation="MDSSSSVKYIKSKSPVKKTLPEESKOPYILSPMDYAMLSTVOYQIDKGLFHKPPLDSDITLLEKLKLSDLAVTLVHPYIPLAGRLSLITERPKSYAFVPCNDSPGAPFIATSDCLIKDIVGAKYPSIVOSFPDHKAHHVDHSHLSLVQYTELDVGDFIGTISNMHAMDGTAFFWKFETAWSEIEFGQCSNONDDICLNPPYLKRY IDEGGPGLPEFLSLPHDEPIRTIESPLIKRMFCFSETIRMTKTRVNOCGTTISSFQSITRAVIYDGCRCITRARIPEDIRETSORVAADNGRWPPLHKRYFGCISALPTAAKAGELIENDGCFPAALKYQVAHEHTSEKYSQMIIDMWLKSPITYIHIDRLPEPKSVMMGSPPRNKYGECEFLGSGVTLRSGYIAHKFDGKVSAIFGRBGGSIDLEVCLVPEFMALBSDEFMKSLVSI"  |
| CDS                       | exon  |
| FEATURES                  | source  |







```

repeat_region      86222..86726
                    /rpt_family="L1"
repeat_region      86811..87050
                    /rpt_family="L1"
repeat_region      complement(92788..93067)
                    /rpt_family="ALU"
BASE COUNT        29577 a 17050 c 18026 g 31564 t
ORIGIN

```

```

Query Match      4.1%; Score 41; DB 9; Length 96217;
Best Local Similarity 47.5%; Pred. No. 1.8;
Matches 122; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

```

```

QY 379 atagctcaacattttttaaaggaagcagatggcaacacatgcatltaacaa 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34580 ATACAGTGAACATAGCTTTATATATCAAAATGATATATTAAACCAAGATTAAAGGAAAA 34521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 439 ggaattatattcaaacacacatgcatgcaattctcaaacatgacccgacgagtc 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34520 TAAAAATATTTCTTATTAACAAAGATTAAGACATTTGTGCTACACAGACGTGTC 34461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 499 atgcgagtggaagcgaagcgaactgaataatcaacatccagtgctcgagtcgagaagagga 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34460 AACGATGTTGAAGAAGAAATACCTTCAGATTGAAGAAAGATACCAATGGAACCTGGA 34401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 559 tgacactgaagatgcgcatcgcattcattacatcatgataacataatgta 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34400 TCAGCTCAAAATGAATGACAGACACTGAAAGGTTAAATATGTGAGCAAAATATAAAGACA 34341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 619 cccataattgtttt 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34340 TTCCATAAATGTCTTT 34324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
LOCUS      AL161624      133021 bp      DNA      linear      PRI 18-MAY-2001
DEFINITION Human DNA sequence from clone RP11-48719 on chromosome X, complete
sequence.
ACCESSION      AL161624
VERSION      AL161624.7      GI:14160916
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
                Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 133021)
AUTHORS      Bird,C.
TITLE      Direct Submission
JOURNAL      Submitted (18-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
                CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
                requests: clonerequests@sanger.ac.uk
                On May 20, 2001 this sequence version replaced gi:13274323.
                During sequence assembly data is compared from overlapping clones.
                Where differences are found these are annotated as variations
                together with a note of the overlapping clone name. Note that the
                variation annotation may not be found in the sequence submission
                corresponding to the overlapping clone, as we submit sequences with
                only a small overlap as described above.
                This sequence was finished as follows unless otherwise noted: all
                regions were either double-stranded or sequenced with an alternate
                chemistry or covered by high quality data (i.e., phred quality >=
                30); an attempt was made to resolve all sequencing problems, such
                as compressions and repeats; all regions were covered by at least
                one plasmid subclone or more than one M13 subclone; and the
                assembly was confirmed by restriction digest. The following
                abbreviations are used to associate primary accession numbers given
                in the feature table with their source databases: Em:, EMBL; Sw:,
                SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
                database can be found at
                http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
                was generated from part of bacterial clone configs of human

```

Chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>  
 RP11-48719 is from the library RP01-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6  
 IMPORTANT: This sequence is not the entire insert of clone RP11-48719. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-212822 is at 132922 in this sequence. The true right end of clone RP11-274W8 is at 100 in this sequence.

#### FEATURES

| source                 | Location/Qualifiers                                     |
|------------------------|---|
| 1..133021              | /organism="Homo sapiens"                                |
| /db_xref="taxon:9606"  |   |
| /chromosome="X"        |   |
| /clone="RP11-48719"    |   |
| /clone_1fb="RP01-11.2" |   |
| 118..425               | /note="AluSg repeat: matches 1..300 of consensus"       |
| 982..1042              | /note="MER69 repeat: matches 1..62 of consensus"        |
| 1311..1622             | /note="AluSg repeat: matches 1..310 of consensus"       |
| 2450..2552             | /note="L2 repeat: matches 2636..2750 of consensus"      |
| 5104..5413             | /note="AluSg repeat: matches 1..304 of consensus"       |
| 5900..6018             | /note="L2 repeat: matches 2486..2568 of consensus"      |
| 6019..6316             | /note="AluSg repeat: matches 1..304 of consensus"       |
| 6317..6514             | /note="L2 repeat: matches 2187..2486 of consensus"      |
| 7589..7901             | /note="AluSg repeat: matches 1..313 of consensus"       |
| 8993..9101             | /note="MIR repeat: matches 156..262 of consensus"       |
| 9102..9441             | /note="Tigerg3b repeat: matches 887..1231 of consensus" |
| 9445..9971             | /note="Tigerg3b repeat: matches 2..540 of consensus"    |
| 9974..10116            | /note="MIR repeat: matches 1..141 of consensus"         |
| 10516..10812           | /note="AluSg repeat: matches 1..297 of consensus"       |
| 10815..11134           | /note="AluSg repeat: matches 1..291 of consensus"       |
| 11810..12051           | /note="MIR repeat: matches 23..262 of consensus"        |
| 13800..14023           | /note="MER53 repeat: matches 1..189 of consensus"       |
| 14038..14282           | /note="MIR repeat: matches 10..262 of consensus"        |
| 14661..14723           | /note="LMB1 repeat: matches 6105..6169 of consensus"    |
| 14724..14834           | /note="AluSg/X repeat: matches 1..111 of consensus"     |
| 14836..14980           | /note="MER20 repeat: matches 69..218 of consensus"      |
| 15747..15910           | /note="MIR repeat: matches 84..255 of consensus"        |
| 17000..17310           | /note="AluSg repeat: matches 1..312 of consensus"       |
| 17611..17646           | /note="MIR repeat: matches 199..234 of consensus"       |
| 17647..17944           | /note="AluSg repeat: matches 2..299 of consensus"       |
| 17945..18142           | /note="MIR repeat: matches 1..199 of consensus"         |
| 18160..18382           |   |





```

Oy      661      ttctacacgaatgcacatagctgcccga 692
Db      63390      TTTTGTGTCACAAACAACATGCTGTCGCA 63421

RESULT   9
AP000736      81580 bp      DNA      linear      PLN 27-DEC-2000
LOCUS      Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone:K17E7.
DEFINITION  AP000736 BA000014
ACCESSION   AP000736.3 GI:8643906
KEYWORDS
SOURCE      Arabidopsis thaliana (strain:Columbia) DNA, clone_11b:Mitsui TAC
            clone:K17E7.
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE   1 (sites)
AUTHORS     Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
TITLE       Structural analysis of Arabidopsis thaliana chromosome 3. II.
            Sequence features of the 4,251,695 bp regions covered by 90 pl, TAC
            and BAC clones
JOURNAL     DNA Res. 7 (3), 217-221 (2000)
MEDLINE     20363099
REFERENCE   2 (bases 1 to 81580)
AUTHORS     Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
TITLE       Direct Submission
JOURNAL     Submitted (18-NOV-1999) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research, 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:yinakam@kazusa.or.jp,
            Tel:81-438-52-3935, Fax:81-438-52-3934)
            On Jun 29, 2000 this sequence version replaced gi:6451951.
COMMENT     Address for correspondence: kaos@kazusa.or.jp
            For the latest information on annotation of this clone, please see
            http://www.kazusa.or.jp/kaos/cgi-bin/sgd.graph.cgi?c=K17E7
            Genes with similarity to proteins in the databases are described in
            'product' or 'note' qualifiers. Genes that have no significant
            protein similarity are described as 'unknown protein'.
            The software programs used to predict genes include: Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://combio.ornl.gov/grail-1.3/),
            GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
            Netemex (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark, http://www.cbs.dtu.dk/services/Netemex2/) and
            SplicePredictor (Volker Brendel, Stanford University,
            http://gremlin1.zool.iastate.edu/cgi-bin/dsp.cgi).
            Genes encoding tRNAs are predicted by tRNAscan-SE
            (Sean Eddy, Washington University School of Medicine, St. Louis,
            http://genome.wustl.edu/eddy/tRNAscan-SE/).
            This sequence may not be the entire insert of this clone. It may be
            shorter because we remove overlaps between neighboring submissions.
            The 5' clone is T26G12 and the 3' clone is T20F20.

FEATURES
    source
        1..81580
            /organism="Arabidopsis thaliana"
            /strain="Columbia"
            /db_xref="taxon:3702"
            /chromosome="3"
            /clone="K17E7"
            /clone_11b="Mitsui TAC"
            /note="contains similarity to retroelement pol
            polyprotein
            gene_id:K17E7.1"
            /codon_start=1
            /pseudo
            /evidence=not_experimental
            join(6774..6898,6993..7357,7439..7782)
            /note="gene_id:K17E7.2"
            /codon_start=1
            /pseudo

CDS
        1..81580
            /evidence=not_experimental
            /product="transposon protein-1ike"
            join(8339..8392,8478..8714,8838..9266)
            /note="contains similarity to En/Spm transposon protein
            gene_id:K17E7.3"
            /codon_start=1
            /evidence=not_experimental
            /protein_id="BAB02991.1"
            /db_xref="GI:13477078"
            /translation="MHKHLAGOKYPMOVLJOEMEBOLKRRVSGFEVFMKTHRVGTFV
            DOKSGHVEAYMKMLERMSPEDKDPQISGNSGSHRMLSIDENKEIFPKLTID
            KGNPFGISLVETLNKRRRRESYASSSTYAOQDLOQLKMSBQAEONACDEHRO
            SOSRLASLEKLVLTFKKDPEDLAVLSDSDIVEVTLPTTHHNSNRHSOPTNHRHS
            RHQRQQRHHHTLSLVSFSYV"
            complement(join(9810..10123,10217..10337))
            /note="emb|CAB88116.1
            gene_id:K17E7.4
            similar to unknown protein"
            /codon_start=1
            /evidence=not_experimental
            /protein_id="BAB02992.1"
            /db_xref="GI:13477079"
            /translation="MLHRSDLGEFAFLGSGNGTKEMKLVSDDAKMDOSQOVSFSDIV
            NLVYDSEIDEDCSYMGKPKRKSIFPKPKRKEPTTITIOENLQFRTPLLVKWLQ
            TRLRLSLDDIKPSLNKRVAKMSPISDELRLRSYSRKATYV"
            join(112471..12477,12881..13057,13190..13277,13381..13486)
            /note="gene_id:K17E7.5
            unknown protein"
            /codon_start=1
            /evidence=not_experimental
            /protein_id="BAB02993.1"
            /db_xref="GI:13477080"
            /translation="MIVAMNVAISFNAGLGSFNSNTTPTSESTITRAVTHRSVSSAI
            EAEALAIALTAQALGCFSEFIPSSANGCEDLAKHALCAFESENSVPRHVKOKSYSG
            VKRSQSGHTHPKSSSANGFL"
            join(144293..14724,14767..15851,15918..16290,16333..17136,
            17203..18798)
            /note="gene_id:K17E7.6"
            /codon_start=1
            /pseudo
            /evidence=not_experimental
            /product="retroelement pol polyprotein-1ike"
            complement(20744..29799)
            /note="contains similarity to TNP2-1ike transposon
            protein
            gene_id:K17E7.7"
            /codon_start=1
            /pseudo
            /evidence=not_experimental
            join(25276..26346,26136..26343,26806..27063,27478..27585,
            27647..27820,28233..28440,29579..29743)
            /note="gene_id:K17E7.8
            unknown protein"
            /codon_start=1
            /evidence=not_experimental
            /protein_id="BAB02994.1"
            /db_xref="GI:13477081"
            /translation="MTYQFPFPCNLNRYGKYSSQNLRSNELVMFPRDLFVGEVEV
            CRDLFVGEVEVCRRDLFVKKFVGGNSNSNLSFIEKNSANLKYSSAKKRFV
            ELYSARKRVEEYISANLRFVGGNSANLSPFGONISANLSFIEKNSANLSS
            FLAYSQTCENIRKVCDDGIASMSCSGEELYSSKRRVRELYSSAKKRFGEELYSSA
            NLRFVGENNISANLSFIRENNLSANLSHISYRRCVACRGTOPAPAKPSQICDVAFAG
            HMLANLEIYSSAKKRFVEIYSSYTRKRFGEIYSSANLRFVGENNSANLSFIEKNS
            SANLSRIYSRRKCVAIYGGTOPQSPKRFGEIYSGIMLSTOFPNCLANLGGKYLSS
            LRVILFPRRDLFVGEKPSLICDVAFAGHDPHVAVSSQLSHFVEIRIFTKATFYRLG
            PVPSSQSRNSVAIT"
            complement(35843..36227)
            /note="contains similarity to non-ITR retroelement reverse
            transcriptase
            gene_id:K17E7.9"
            /codon_start=1
            /pseudo
            /evidence=not_experimental

```

QY 488 ccgacgagtcacatcgaggtggaacagaaagacatcaaatcaacatc 534  
 11 1111 1 1 11111111111111111111  
 Db 42546 TTGATATAGTCAGAACCAATGTAACCGAATTAAGTAATATCCACTTC 42500  
 RESULT 10  
 AC027033 98017 bp DNA linear PLN 11-OCT-2000  
 LOCUS Arabidopsis thaliana chromosome 1 BAC F21N10 genomic sequence,  
 AC027033/c complete sequence.  
 DEFINITION  
 AC027033  
 AC027033 GI:10092252  
 KEYWORDS  
 SOURCE HTG.  
 ORGANISM  
 Arabidopsis thaliana  
 thale cress.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 98017)  
 AUTHORS Lin,X., Kaul,S., Town,C.D., Bentto,M.-T., Greasy,T.H., Haas,B.J.,  
 Wu,D., Malti,R., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R.,  
 Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.  
 Arabidopsis thaliana chromosome 1 BAC F21N10 genomic sequence  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 98017)  
 REFERENCE Lin,X. and Kaul,S.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (26-MAR-2000) The Institute for Genomic Research, 9712  
 TITLE Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org  
 REFERENCE 3 (bases 1 to 98017)  
 AUTHORS Town,C.D. and Kaul,S.  
 JOURNAL Direct Submission  
 TITLE Submitted (12-SEP-2000) The Institute for Genomic Research, 9712  
 JOURNAL Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org  
 COMMENT On Sep 12, 2000 this sequence version replaced gi:7709919.  
 Address all correspondence to:at@tigr.org  
  
 BAC clone F21N10 is from Arabidopsis thaliana chromosome 1  
 The orientation of the sequence is from S56 to T7 end of the BAC  
 clone.  
 Genes were identified by a combination of several methods: Gene  
 prediction programs including Genscan<sup>†</sup> (Chris Burge,  
 http://CCR-081.mit.edu/GENSCAN.html), Genemark<sup>†</sup> (Mark Borodovsky,  
 http://genemark.biology.gatech.edu/Genemark/), Glimmer<sup>†</sup> (a variant  
 of Glimmer<sup>†</sup>, see Mihaela Perlea,  
 http://www.tigr.org/softlab/glimmer.htm/glimmer.html, and  
 GeneSplicer (Mihaela Perlea and Steven Salzberg, contact  
 mperlea@tigr.org), searches of the complete sequence against a  
 peptide database and the plant EST database at TIGR  
 (http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to  
 indicate the level of evidence for their annotation. Genes with  
 similarity to other proteins are named after the database hits.  
 Genes without significant peptide similarity but with EST  
 similarity are named as unknown proteins. Genes with protein  
 or EST similarity, that are predicted by more than two gene  
 prediction programs over most of their length are annotated as  
 hypothetical proteins. Genes encoding tRNAs are predicted by  
 tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).  
 Simple repeats are identified by repeatmasker (Arian Smit,  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html).  
 Location/Qualifiers  
 1..98017  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="F21N10"  
 1184..5698  
 /gene="F21N10.12"  
 /note="pseudogene, polyprotein"  
 /pseudo



[illegible]

SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:BJ158C05.  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 65955)  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-MAY-2001) Takuji Sasaki, National Institute of  
 Agrobiological Sciences, Rice Genome Research Program, Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://ryp.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 COMMENT On Feb 12, 2002 this sequence version replaced gi:14192851.  
 The orientation of the sequence is from M13rev to -21M13 of the BAC  
 clone.  
 FEATURES  
 source Location/Qualifiers  
 1..65955  
 /organism="Oryza sativa"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /chromosome="1"  
 /clone="BJ158C05"  
 BASE COUNT 19002 a 13818 c 13940 g 19195 t  
 ORIGIN  
 Query Match 4.0%; Score 40; DB 8; Length 65955;  
 Best Local Similarity 50.8%; Pred. No. 3.3;  
 Matches 121; Conservative 0; Mismatches 115; Indels 2; Gaps 1;  
 QY 249 cgtgctgcgtgacacatgagggtcatctatgaggcgtcgaggaagaagagagcgcaca 308  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 51338 CGCCGTGTCGCCGAACACAGAAGGCGATGTCGGATCGTAACCCCTGGATGTGACATATTC 51269  
 QY 309 tgaacacatgatgtgtctatcagccgcatgagaggaagcaacatgacctttccct 368  
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 51268 AGTATACATGATGATTTAGATAGAGGTATGTCCAAATTCGTATCTACTAGAAATATGTCATAT 51209  
 QY 369 ctggaatctatagctcacacttttttcaatgagaagaagatgtgcaaacacatgcat 428  
 || || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 51208 CT-AATTACATAGATGTGGTTTATTATCGACGAGGAGATCAATCACACACACTTGCAT 51151  
 QY 423 ttcaacaagaagaattaatcttcaaacaccatgacatgcatcttcaaacatgc 486  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 51150 TTTCGGACACGAAAGACTTAACCTTAAGCGCTGCAATGCTTCATGCTTACTCTGTGC 51093  
 RESULT 13  
 LOCUS PFMAL13P3  
 DEFINITION Plasmidium falciparum chromosome 13 strain 3D7, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, In unordered pieces.  
 ACCESSION AL049184.5 GI:5763803  
 VERSION HTG: HTGS-PHASE1.  
 KEYWORDS malaria parasite P. falciparum.  
 SOURCE Plasmidium falciparum  
 ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.  
 REFERENCE 1 (bases 1 to 318221)  
 AUTHORS Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M.  
 and Barrell,B.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-MAR-1999) P.falciparum Genome Sequencing Consortium,  
 The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge  
 CB10 1SA, UK  
 COMMENT On Aug 24, 1999 this sequence version replaced gi:5731882.  
 For more information about this sequence or the Malaria Project,  
 see http://www.sanger.ac.uk/Projects/P.falciparum. IMPORTANT: This  
 sequence is unfinished and does not necessarily represent the  
 correct sequence. Work on the sequence is in progress and the  
 release of this data is based on the understanding that the  
 sequence may change as work continues. The sequence may be





Search completed: September 12, 2002, 07:27:09  
Job time: 20617 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 04:57:23 ; Search time 406.92 Seconds  
(without alignments)  
4219.292 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_1\_1000  
Perfect score: 1000  
Sequence: 1 agaacaccctcatttaga.....acactccccgtgggtccct 1000

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|-------------|
| 1          | 1000  | 100.0       | 11473     | 20 | AAK34650    |
| 2          | 1000  | 100.0       | 11475     | 22 | AAH78338    |
| 3          | 39.4  | 3.9         | 4590      | 22 | AAH24065    |
| 4          | 36    | 3.6         | 15548     | 22 | ABL34155    |
| 5          | 35.6  | 3.6         | 1741      | 18 | AAH84300    |
| 6          | 35.6  | 3.6         | 1741      | 19 | AAH52578    |
| 7          | 35.2  | 3.5         | 8387      | 24 | ABL32247    |
| 8          | 34.6  | 3.5         | 710       | 22 | AAI97878    |
| 9          | 34.6  | 3.5         | 32220     | 22 | AAI37568    |

|    |      |     |        |    |          |                      |
|----|------|-----|--------|----|----------|----------------------|
| 10 | 34.4 | 3.4 | 362    | 22 | AAK57698 | Human immune/haema   |
| 11 | 34.4 | 3.4 | 9725   | 24 | ABL33293 | Human immune syste   |
| 12 | 34.4 | 3.4 | 20394  | 22 | AAH24892 | Pimaricin biosynth   |
| 13 | 34.2 | 3.4 | 6754   | 24 | AAH61304 | Human gene regulat   |
| 14 | 34.2 | 3.4 | 12767  | 22 | ABH17068 | Human nervous syst   |
| 15 | 34.2 | 3.4 | 101786 | 21 | AAH22293 | BAC containing rep   |
| 16 | 34.2 | 3.4 | 350    | 16 | AAO99380 | Rat allograft infl   |
| 17 | 34   | 3.4 | 350    | 21 | AAH65600 | Rat allograft infl   |
| 18 | 34   | 3.4 | 1279   | 23 | ABL07703 | Drosophila melanog   |
| 19 | 34   | 3.4 | 3640   | 23 | ABL07702 | Drosophila melanog   |
| 20 | 34   | 3.4 | 5340   | 22 | AAH26289 | P. falciparum egf    |
| 21 | 34   | 3.4 | 7810   | 22 | AAH45434 | Chemically pretrea   |
| 22 | 34   | 3.4 | 56153  | 22 | AAH46794 | Tumour suppressor    |
| 23 | 33.8 | 3.4 | 6415   | 22 | AAH46371 | Tumour immune syste  |
| 24 | 33.8 | 3.4 | 4590   | 22 | AAH24065 | Yeast AOD9604-asso   |
| 25 | 33.6 | 3.4 | 628    | 21 | AAH40041 | Arabidopsis thalia   |
| 26 | 33.4 | 3.3 | 2324   | 22 | AAH16057 | Human cdna sequenc   |
| 27 | 33.4 | 3.3 | 1000   | 21 | AAH02484 | Human colon cancer   |
| 28 | 33.2 | 3.3 | 5886   | 24 | ABL34212 | Human immune syste   |
| 29 | 33.2 | 3.3 | 11812  | 22 | AAH45501 | Chemically pretrea   |
| 30 | 33.2 | 3.3 | 11812  | 22 | AAH46741 | Tumour suppressor    |
| 31 | 33.2 | 3.3 | 11812  | 22 | ABL34118 | Human immune syste   |
| 32 | 33.2 | 3.3 | 16771  | 22 | AAH76778 | Human immune/haema   |
| 33 | 33.2 | 3.3 | 17763  | 23 | ABL02400 | Drosophila melanog   |
| 34 | 33.2 | 3.3 | 168575 | 22 | AAH21613 | Human hypocrerin r   |
| 35 | 33.2 | 3.3 | 495    | 21 | AAH36252 | Arabidopsis thalia   |
| 36 | 33   | 3.3 | 6297   | 24 | ABL32995 | Human immune syste   |
| 37 | 33   | 3.3 | 15923  | 24 | AAH61132 | Human gene regulat   |
| 38 | 33   | 3.3 | 32176  | 22 | AAH36453 | Human cardiovascular |
| 39 | 33   | 3.3 | 32178  | 22 | AAH36454 | Human cardiovascular |
| 40 | 33   | 3.3 | 347    | 22 | AAH70921 | Human immune/haema   |
| 41 | 32.8 | 3.3 | 347    | 22 | AAH70922 | Human immune/haema   |
| 42 | 32.8 | 3.3 | 700    | 22 | AAH24329 | Human breast cance   |
| 43 | 32.8 | 3.3 | 4329   | 22 | AAH30492 | DNA encoding novel   |
| 44 | 32.8 | 3.3 | 4329   | 22 | AAH30492 | DNA encoding novel   |
| 45 | 32.8 | 3.3 | 4329   | 22 | AAH06272 | Human reproductive   |

## ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 |  |
| ID       | AAK34650 standard; DNA: 11473 BP.                                      |
| XX       | AAK34650:  |
| XX       | 05-JUL-1999 (first entry)  |
| DE       | Starch branching enzyme II (SBE II) gene sequence.                     |
| XX       |  |
| XX       | Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBS; |
| KW       | starch branching enzyme; starch soluble synthase; debranching enzyme;  |
| KW       | endosperm; wheat; barley; granule-bound synthase; glutenin; starch;    |
| KW       | grain softness protein I; bacterial isomylase; glycogen synthase;      |
| KW       | WBE I-D4 gene; ss.   |
| XX       |  |
| OS       | Triticum tauschii.   |
| XX       |  |
| PN       | MO9914314-AI.  |
| XX       |  |
| PD       | 25-MAR-1999.   |
| XX       |  |
| PF       | 11-SEP-1998; 98MO-AU00743.   |
| XX       |  |
| PR       | 20-MAR-1998; 98AU-0002509.   |
| XX       |  |
| PR       | 12-SEP-1997; 97AU-0009108.   |
| XX       |  |
| PA       | (CSIR ) COMMONWEALTH SCI & IND RES ORG.                                |
| PA       | (GOOD-) GOODMAN FIELDER LTD.   |
| PA       | (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.                                 |
| PA       | (AUSU ) UNIV AUSTRALIAN NAT.   |
| XX       |  |





06-NOV-2000; 2000MO-AU01362.  
05-NOV-1999; 99AU-0003875.  
(META-) METABOLIC PHARM LTD.  
Belyea CI, Ng FM, Vaughan P;  
WPI; 2001-328876/34.  
New organisms containing nucleic acid encoding a growth hormone  
fragment which modulates lipid metabolism are useful to produce dietary  
aids for obesity and in the meat production industry -  
Disclosure; Page 48-50; 54pp; English.  
The invention relates to novel transgenic organisms useful in the  
production of functional food and drink products for the treatment  
or prevention of obesity via the regulation of lipid metabolism. The  
organisms comprise a polynucleotide encoding a growth hormone fragment  
capable of stimulating the activity of hormone-sensitive lipase (the key  
enzyme in lipolysis) and inhibiting acetyl CoA carboxylase (the key  
enzyme in lipogenesis). The growth hormone fragment preferably contains  
at least the disulphide-bonded loop of a mammalian growth hormone (but is  
not the full-length growth hormone) and is optionally linked to an  
epitope tag or heterologous fusion protein partner. The transgenic  
organism may be a microorganism used to produce a fermented product  
(e.g., yeast), or an edible plant or animal or cell thereof. Food or  
drink made using methods of the invention are used to modify fat/lean  
ratio, lipid metabolism or food use in a mammal. In particular, the food  
or drink products may be used to treat or prevent obesity, particularly  
in humans, and may also be used to improve the fat/lean ration of  
livestock raised for meat production. In the exemplification of the  
invention, the human growth hormone (hGH) fragment analogue AOD9604 was  
expressed in yeast, optionally fused to the FLAG epitope (AAB33625).  
The present sequence is described as a DNA sequence from yeast in  
the sequence listing, but is not further referred to in the  
specification.

[illegible][illegible]

RESULT 4  
ABL34155/c  
ID ABL34155 standard; DNA; 15548 BP

AC ABL34155;

DT 26-MAR-2002 (first entry)

Human immune system associated gene SEQ ID NO: 2128.

Human; immune system disease; cytosine methylation; antiasthmatic; KW

KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

KW antiinflammatory; cancer; eye disease; arteriosclerosis; and

KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;

XX

XX

XX

XX

XX

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

**PT** Nucleic acid comprising fragment of chemically modified gene, useful

PT cytosine methylation -

PS Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated

CC can be used in the diagnosis and treatment of immune system disorders,

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

XX

| Query Match | 3.68; | Score 36; | DB 24; | Length 15548; |
|-------------|-------|-----------|--------|---------------|
|-------------|-------|-----------|--------|---------------|

|         |     |              |    |            |     |        |    |      |    |
|---------|-----|--------------|----|------------|-----|--------|----|------|----|
| Matches | 93; | Conservative | 0; | Mismatches | 95; | Indels | 0; | Gaps | 0; |
|---------|-----|--------------|----|------------|-----|--------|----|------|----|

427 atttcaacaagaaaattattctcaaacaccatgacatgcaattctcaaacatgc 486

Db 7849 ATCTAAATCAACTTAAAAAATTATTTTACCAAATTAATAATACATTTCATTAACACACAC 7790

QY 487 accgacgagtcctatgcgaggtggaacgaagaactgaaatcaacatcccgattgtcgag 546

```
Db 7789 CTTAAAAAACTTACACATATACCCAAATATCAAAATCCAACTTACTTTATACAT 7730
QY 547 tcgagaagagatgacactgaagatgctgctatcattcatcttaccatagatgacaa 606
Db 7729 TTTAAAAAAACATATCATATACATCTTAAATATACATTTTAAATTCGATCTCAAAA 7670
QY 607 atacataa 614
Db 7669 AAACAAAA 7662
```

## RESULT 5

```
AA84300
ID AAT84300 standard; cDNA: 1741 BP.
XX
AC AAT84300;
XX
DT 10-NOV-1997 (first entry)
XX
DE Quail Flt4 receptor tyrosine kinase ligand VEGF-C cDNA.
XX
KW VEGF-C; Flt4; receptor tyrosine kinase; VEGFR-3; quail;
KM vascular endothelial growth factor receptor-3; ligand;
XX angiogenesis; wound healing; lymph vessel; therapy; diagnosis; ss.
OS Coturnix coturnix.
```

```
XX
FH Key Location/Qualifiers
FT CDS 453..1709
FT /tag= a
XX
```

WO9705250-A2.

13-FEB-1997.

01-AUG-1996; 96WO-F100427.

28-JUN-1996; 96US-0671573.

01-AUG-1995; 95US-0510133.

12-JAN-1996; 96US-0585895.

14-FEB-1996; 96US-0601132.

(UYHE-) UNIV HELSINKI LICENSING LTD OY.

Alitalo K, Joukov V;

WPI; 1997-145688/13.

P-PSDB; AAM00934.

Example 20; Page 124-126; 183pp; English.

This cDNA clone codes for quail VEGF-C (AAM00934), a novel ligand that binds specifically to the Flt4 receptor tyrosine kinase (VEGFR-3), stimulating phosphorylation of the receptor. The clone was isolated from a quail cDNA library using mouse (see AAT84277) CC and human (see AAT84276) VEGF-C cDNA fragments as probes. The CC isolated polynucleotide can be used to produce recombinant CC polypeptides corresponding to non-human mammalian variants of VEGF-C.

Sequence 1741 BP; 445 A; 455 C; 449 G; 392 T; 0 other;

Query Match 3.6%; Score 35.6; DB 18; Length 1741;  
Best Local Similarity 55.7%; Pred. No. 1.9; Mismatches 0; Gaps 0;  
Matches 68; Conservative 0; Indels 0;

QY 825 cgcgcgtgcgcgtgtaacgcgcctccgcgcgtgtgttcgcgtctcctccgcgcgc 884

```
Db 41 cgcgccgcggagggcgcgcgtgcgcgcgcacactgggctcctcctcctcctcc 100
QY 885 agcgtccctccacgcgtccgcgtcgcgcacccctcgtgcgcgcgcacgaaggga 944
Db 101 tctcctcctccctcctcctcctcgtcgcgttcacacgcgcgcgcgcgcgcgcga 160
QY 945 gg 946
Db 161 tg 162
```

## RESULT 6

```
AAV52578
ID AAV52578 standard; cDNA: 1741 BP.
XX
AC AAV52578;
XX
DT 14-DEC-1998 (first entry)
XX
DE Quail vascular endothelial growth factor C gene.
XX
KW Flt4; vascular endothelial growth factor C; vascular endothelial cell;
KM lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
XX lymphangiogenesis; oedema; elephantiasis; Milroy's disease; ss.
OS Coturnix sp.
```

```
XX
FH Key Location/Qualifiers
FT CDS 453..1709
FT /tag= a
FT /product= "VEGF-C"
XX
```

WO9833917-A1.

06-AUG-1998.

02-FEB-1998; 98WO-US01973.

05-FEB-1997; 97US-0795430.

(LUDW-) LUDWIG INST CANCER RES.

(UYHE-) UNIV HELSINKI LICENSING LTD.

Alitalo K, Joukov V;

WPI; 1998-437470/37.

P-PSDB; AAM75743.

New isolated vascular endothelial growth factor polypeptide(s) - used to develop products for treating, e.g. cancers, inflammation, oedema, granulocytopenia or for wound healing or tissue transplantation

Example 20; Page 121-123; 177pp; English.

The vascular endothelial growth factor C (VEGF-C) polypeptides have activities affecting growth and migration of vascular endothelial cells, promoting growth of lymphatic endothelial cells and lymphatic vessels, increasing vascular permeability, and affecting myelopoiesis. The CC products can be used for stimulating angiogenesis, for inhibiting angiogenesis, for stimulating lymphangiogenesis, treatment or prevention of inflammation, oedema, elephantiasis, or Milroy's disease. They can also be used to modulate myelopoiesis, e.g. treating granulocytopenia. They can also be used for modulating the growth of endothelial cells. They can also be used to stimulate lymphocyte production and maturation, and to promote or inhibit trafficking of leucocytes between tissues and lymphatic vessels or to affect migration in and out of the thymus.

Sequence 1741 BP; 445 A; 455 C; 449 G; 392 T; 0 other;

Query Match 3.6%; Score 35.6; DB 19; Length 1741;





Db 329 ATCAGAGCTTCAACCCAGGCTTTATAGCAAAAAAAAAAATAATCAAAAACCCCTTAG 270  
Qy 464 gacatgcattctcaaacatgacgcgagtcctcatgctgaggtggaacgaagaactga 523  
Db 269 GACTTCATCACTCTAGCCCTGACCTTCACAGTCCTCTGTGAGGGGCTCTCAAGACCCCTT 210  
Qy 524 aatcacatcccggtgtgctgagtcgagaagaagatgacactgaagta 572  
Db 209 AAAATAATATTACATAGGACTTGTGACTTCATTTCACAGTGGAATA 161

RESULT 9  
ID AAL37568 standard; DNA: 32220 BP.  
XX AAL37568;  
XX AAL37568;  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3933.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW candidant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PE 17-JAN-2001; 2001WO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236337.  
PR 29-SEP-2000; 2000US-0236357.  
PR 29-SEP-2000; 2000US-0236358.  
PR 29-SEP-2000; 2000US-0236359.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.



```

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

```

```

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250161.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-483426/52.
DR P-PSDB; AAM84917.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Claim 1; SEQ ID NO 2758; 3071pp + Sequence Listing; English.
XX
XX AAK4951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 362 BP; 89 A; 80 C; 85 G; 106 T; 2 other;

Query Match 3.4%; Score 34.4; DB 22; Length 362;
Best Local Similarity 57.4%; Pred. No. 1.9; Mismatches 46; Indels 0; Gaps 0;
Matches 62; Conservative 0;

QY 768 taaatgttttgaacaataactataaagaggtactagagcgccgaagc 827
|| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 tatagttttttaacttaactaataattttcacaagtcagtcagtcgaatcc 302
|| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 828 catgcccaggtaaagcgctccagcgctgtgtgttcgcatctgctct 875
| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 aagatcaagtgccagcagcagtcagtttgytaaggtctcttct 350

RESULT 11
ABL33293/c
ID ABL33293 standard; DNA; 9725 BP.
XX
AC ABL33293;
XX

```

DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 1266.  
XX  
XX Human: immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytosolic; noctropic;  
KW antiprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antineuritic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200200928-A2.  
XX  
XX 03-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
XX 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPIC-) EPIDEMIOLOGICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI: 2002-130909/17.  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
XX  
XX  
PS Claim 1; SEQ ID NO 1266; 32pp + Sequence listing; German.  
XX  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX  
XX Sequence 9725 BP; 2521 A; 335 C; 2382 G; 4487 T; 0 other;  
SQ

Query Match 3.4%; Score 34.4; DB 24; Length 9725;  
Best Local Similarity 48.5%; Pred. No. 13;  
Matches 95; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
OY 305 cacatgaacacacatgatctatcagcctgatgagggaggaacacatgacctttc 364  
DB 7136 CTCCTATACCTCTTAACTTAAATTAACCTATTAAATTAATCACCACCAAAATCATC 7137  
OY 365 ccccttgaattcatgacacacttttttaatggaagaagatgagcaacacat 424  
DB 7136 TCAATTAATAATTAATTAATCTCTCATTTTAAAAATAAAAAAGCCTAATTAATAAAA 7077  
OY 425 gcaatttcaacaagaataatcctcaaacacacacatgacatgcaattcacaacat 484  
DB 7076 ACAATTTAAATACCAAAATCATCTCGCAAAAAAATAATTAATTAACGTTCTTAAAAAC 7017  
OY 485 gcaacgacgagtcacat 500  
DB 7016 GCAACCAAAATTAAT 7001

RESULT 12  
AAF24892  
ID AAF24892 standard; DNA: 20394 BP.  
XX

AC AAF24892;  
XX  
XX 20-APR-2001 (first entry)  
XX  
XX Pimaricin biosynthesis associated polyketide synthase gene.  
DE  
XX  
XX Polyketide synthase; oxidative modification; metabolite; antibiotic;  
KW anticancer; pimaricin; ss.  
XX  
XX Streptomyces natalensis.  
XX  
XX  
XX Key Location/Qualifiers  
FT CDS 1..20394  
FT /tag= a  
FT /product= "polyketide synthase"  
XX  
XX WO200077222-A1.  
XX  
XX 21-DEC-2000.  
XX  
XX 14-JUN-2000; 2000WO-EP06227.  
XX  
XX 14-JUN-1999; 99EP-0201893.  
XX  
XX (STAM ) DSM NV.  
XX  
XX Martin JF, Aparicio JF, Collina AJ;  
XX  
XX WPI: 2001-080693/09.  
XX P-PSDB: AAB31558.  
XX  
XX New polynucleotides encoding enzymes involved in the biosynthesis of  
PT pimaricin, useful for modifying the biosynthesis of pimaricin and in  
PT the synthesis of new compounds -  
XX  
XX  
PS Disclosure: Page 53-80; 116pp; English.  
XX  
XX  
XX The present sequence encodes a polyketide synthase which is associated  
CC with the biosynthesis of pimaricin. The polyketide synthase polypeptide  
CC is useful for the oxidative modification of a methyl group of a suitable  
CC compound, e.g. a bioactive compound including a secondary metabolite,  
CC antibiotics and anticancer agents. Recombinant cells comprising the  
CC gene are useful for the production of pimaricin. The polyketide synthase  
CC polynucleotide may be over expressed in Streptomyces, leading to an  
CC increase in the biosynthesis of pimaricin, as a source of primers for  
CC amplification reaction and as probes.  
XX  
XX Sequence 20394 BP; 2611 A; 7486 C; 7397 G; 2900 T; 0 other;  
SQ

Query Match 3.4%; Score 34.4; DB 22; Length 20394;  
Best Local Similarity 53.8%; Pred. No. 20;  
Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
OY 844 cgtcccaagccgttggttgatctgctcccccgaagcagcgtgcctccaccgtcc 903  
DB 7653 cgcgtccgcgcgcgcgcacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 7712  
OY 904 gtcgtcgtccacactctgctgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 963  
DB 7713 gtgggtcagaagc 7772  
OY 964 acacacacacac 975  
DB 7773 cgacttccacac 7784

RESULT 13  
AAS61304/C  
ID AAS61304 standard; DNA: 6754 BP.  
XX  
XX AAS61304;  
AC  
XX



PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228287.  
PR 01-SEP-2000; 2000US-0228343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0234400.  
PR 14-SEP-2000; 2000US-0234401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI: 2001-541565/60.  
DR  
XX  
XX  
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating nervous system  
PT cancers and metastases -  
XX  
XX  
PS Disclosure; SEQ ID NO 9399; 1701pp + Sequence Listing; English.  
PS  
XX  
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins  
CC (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 12767 BP; 4099 A; 2208 C; 2513 G; 3947 T; 0 other;

Query Match 3.4%; Score 34.2; DB 22; Length 12767;  
Best Local Similarity 54.3%; Pred. No. 18;  
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
  
Oy 717 ggagacgacgacacatcagacactcaactgtcttgcctggagacacataatgctt 776  
||||| | | | | ||||| | | ||||| ||| |  
Db 1381 ggagactgctggagacgaagattcaccactgcactcagcctgggtgacagacatgact 1440







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 04:05:43 ; Search time 92.79 Seconds  
(without alignments)  
2647.198 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_1\_1000

Sequence: 1 agaacaccccccatttaga.....acactcccggtggtccct 1000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/pdata/1/lna/5A.COMB.seq:\*  
2: /cgn2\_6/pdata/1/lna/5B.COMB.seq:\*  
3: /cgn2\_6/pdata/1/lna/6A.COMB.seq:\*  
4: /cgn2\_6/pdata/1/lna/6B.COMB.seq:\*  
5: /cgn2\_6/pdata/1/lna/PCRTUS.COMB.seq:\*  
6: /cgn2\_6/pdata/1/lna/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 39.8  | 4.0         | 7218   | 1  | US-08-232-463-14   |
| 2          | 35.6  | 3.6         | 1741   | 3  | US-08-795-430-12   |
| 3          | 34    | 3.4         | 350    | 1  | US-08-171-385-14   |
| 4          | 34    | 3.4         | 350    | 3  | US-08-361-441B-14  |
| 5          | 32.6  | 3.3         | 15144  | 4  | US-08-458-434A-6   |
| 6          | 32.2  | 3.2         | 1208   | 4  | US-08-979-608A-16  |
| 7          | 31.8  | 3.2         | 35060  | 3  | US-08-814-095-7    |
| 8          | 31.2  | 3.1         | 726    | 4  | US-08-998-416-1085 |
| 9          | 31.2  | 3.1         | 1767   | 1  | US-08-399-646-1    |
| 10         | 31.2  | 3.1         | 1767   | 1  | US-08-607-321-1    |
| 11         | 31.2  | 3.1         | 1767   | 2  | US-08-961-240-1    |
| 12         | 31.2  | 3.1         | 1767   | 2  | US-08-605-501-1    |
| 13         | 31.2  | 3.1         | 2161   | 1  | US-08-399-646-11   |
| 14         | 31.2  | 3.1         | 2161   | 1  | US-08-607-321-11   |
| 15         | 31.2  | 3.1         | 2161   | 2  | US-08-961-240-11   |
| 16         | 31.2  | 3.1         | 2161   | 2  | US-08-605-501-11   |
| 17         | 30.8  | 3.1         | 3284   | 6  | 5258288-3          |
| 18         | 30.6  | 3.1         | 19124  | 2  | US-08-487-826B-13  |
| 19         | 30.4  | 3.0         | 4252   | 2  | US-08-475-844-4    |
| 20         | 30.4  | 3.0         | 4252   | 5  | PCR-US95-08429-4   |
| 21         | 30.2  | 3.0         | 1623   | 4  | US-08-339-214-25   |
| 22         | 30.2  | 3.0         | 7791   | 2  | US-08-149-097D-23  |
| 23         | 30.2  | 3.0         | 7791   | 3  | US-08-949-386-23   |
| 24         | 30.2  | 3.0         | 7791   | 3  | US-08-450-562-23   |
| 25         | 30.2  | 3.0         | 7791   | 4  | US-08-984-709A-23  |
| 26         | 30.2  | 3.0         | 7808   | 4  | US-08-149-097D-22  |
| 27         | 30.2  | 3.0         | 7808   | 3  | US-08-949-386-22   |

|      |      |     |       |   |                    |                    |
|------|------|-----|-------|---|--------------------|--------------------|
| c 28 | 30.2 | 3.0 | 7808  | 3 | US-08-450-562-22   | Sequence 22, Appl  |
| c 29 | 30.2 | 3.0 | 7808  | 4 | US-08-984-709A-22  | Sequence 22, Appl  |
| c 30 | 30   | 3.0 | 3719  | 1 | US-08-920-812-10   | Sequence 10, Appl  |
| c 31 | 30   | 3.0 | 3719  | 1 | US-08-920-827-10   | Sequence 10, Appl  |
| c 32 | 30   | 3.0 | 3719  | 1 | US-08-921-177-10   | Sequence 10, Appl  |
| c 33 | 30   | 3.0 | 3719  | 1 | US-08-362-577C-10  | Sequence 10, Appl  |
| c 34 | 30   | 3.0 | 3719  | 2 | US-08-920-828-10   | Sequence 10, Appl  |
| c 35 | 29.8 | 3.0 | 7676  | 1 | US-08-451-777A-7   | Sequence 7, Appl   |
| c 36 | 29.8 | 3.0 | 7676  | 2 | US-08-451-778A-7   | Sequence 7, Appl   |
| c 37 | 29.8 | 3.0 | 7676  | 2 | US-08-998-208-7    | Sequence 7, Appl   |
| c 38 | 29.8 | 3.0 | 7676  | 5 | PCR-US95-06743-7   | Sequence 7, Appl   |
| c 39 | 29.4 | 2.9 | 976   | 1 | US-08-125-628-7    | Sequence 7, Appl   |
| c 40 | 29.4 | 2.9 | 29598 | 4 | US-09-341-587-6    | Sequence 6, Appl   |
| c 41 | 29.2 | 2.9 | 340   | 1 | US-08-171-385-27   | Sequence 27, Appl  |
| c 42 | 29.2 | 2.9 | 340   | 1 | US-08-361-441B-27  | Sequence 27, Appl  |
| c 43 | 29.2 | 2.9 | 3396  | 4 | US-08-974-549A-640 | Sequence 640, Appl |
| c 44 | 29.2 | 2.9 | 40352 | 3 | US-08-846-111D-15  | Sequence 15, Appl  |
| c 45 | 29   | 2.9 | 1678  | 1 | US-08-261-677-10   | Sequence 10, Appl  |

#### ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMUN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZgpt-F1s  
US-08-232-463-14

```
Query Match      4.0%, Score 39.8; DB 1; Length 7218;
Best Local Similarity 5.0%; Pred. No. 0.023;
Matches 20; Conservative 206; Mismatches 173; Indels 0; Gaps 0;

OY 73 ttcttaaaagacagccatttagccctcttacaagaagctcaacagctccaac 132
    ||| | : : : : : : : : : : : : : : : : : : : : : : : :
DB 1442 tttggtacrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr 1383

OY 133 gctctagatcacccagctgcaagtttagcgagcaccacacagggcgatcga 192
    : : : : : : : : : : : : : : : : : : : : : : : :
DB 1382 rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr 1323

OY 193 actgacagacgctcagcagagagccagcacacagcttagcctcagcagctg 252
    : : : : : : : : : : : : : : : : : : : : : : : :
DB 1322 rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr 1263

OY 253 agtgcgtacacatgggtcatctatggcgctcgagcaaggaagagagcagcatgaa 312
    : : : : : : : : : : : : : : : : : : : : : : : :
DB 1262 rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr 1203

OY 313 caccatgatgctcatcagcgctgagagagcaacacatgccttccctctg 372
    : : : : : : : : : : : : : : : : : : : : : : : :
DB 1202 rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr 1143

OY 373 aaatcatagctcacacttttttaatggaagcaagatgagcaaacatgcatlcc 432
    : : : : : : : : : : : : : : : : : : : : : : : :
DB 1142 rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr 1083

OY 433 aaacaaggaataatctcacaaccacatgacatgca 471
    : : : : : : : : : : : : : : : : : : : : : : : :
DB 1082 rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr 1044

RESULT 2
US-08-795-430-12
; Sequence 12, Application US/08795430
; Patent No. 6130071
; GENERAL INFORMATION:
; APPLICANT: Altalo, Karl
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795.430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/F196/00427
; FILING DATE: 01-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/671,573
; FILING DATE: 28-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,895
; FILING DATE: 12-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/510,133
```

```
; FILING DATE: 01-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/33691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
;
; TELEFAX: 25-3856
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1741 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 453..1706
;
US-08-795-430-12

Query Match      3.6%, Score 35.6; DB 3; Length 1741;
Best Local Similarity 55.7%; Pred. No. 0.22;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

OY 825 cggcatgcccaggaataacgcgcgtccacagcgttggttgcatctcgtctccgacgc 884
    ||| | | | | | | | | | | | | | | | | | | | | | | |
DB 41 cggccggcgagggcgccgctgcgacggcgacacactggctccttcccttcccttcc 100

OY 885 agcgctgcctcacccgtccgcgtcgtcgctgcaacctgctgtgagcgagcaagaagga 944
    ||| | | | | | | | | | | | | | | | | | | | | | | |
DB 101 tctctctctctctctcttctcttctgcttccacgctccgagcgacgacgctcgca 160

OY 945 gg 946
DB 161 TG 162

RESULT 3
US-08-171-385-14/c
; Sequence 14, Application US/08171385
; Patent No. 5527884
; GENERAL INFORMATION:
; APPLICANT: Mary E. Russell
; APPLICANT: Ulrike Utans
; TITLE OF INVENTION: Mediators of Chronic Allograft
; TITLE OF INVENTION: Rejection
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555x
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/171,385
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
```

```

1      REGISTRATION NUMBER: 34,4819
2      REFERENCE/DOCKET NUMBER: 05433/0060001
3      TELECOMMUNICATION INFORMATION:
4      TELEPHONE: (617) 542-5070
5      TELEFAX: (617) 542-8906
6      TELEX: 200154
7      INFORMATION FOR SEQ ID NO: 14:
8      SEQUENCE CHARACTERISTICS:
9          LENGTH: 350
10         TYPE: nucleic acid
11         STRANDEDNESS: double
12         TOPOLOGY: linear
13
14     US-08-171-385-14

```

|                       |                |                |          |            |
|-----------------------|----------------|----------------|----------|------------|
| Query Match           | 3.4%           | Score 34       | DB 1     | Length 350 |
| Best Local Similarity | 46.3%          | Pred. No.      | 0.27     |            |
| Matches 112           | Conservative 0 | Mismatches 130 | Indels 0 | Gaps 0     |

|    |     |  |     |
|----|-----|--|-----|
| QY | 561 | acactgaagaatgcatgcatatcagcttccattcaatacatcgtacaatatgaatgacc    | 620 |
|    |     |  |     |
| Db | 246 | ACAATTATACCCACACAGATTTTCATAGCCATATACATGTATGTCTTACATTAACATTAAGG | 187 |
| QY | 621 | ctacaattgcttcttctggagacagatgctgctcttcttcttcaacgaaatgcac        | 680 |
|    |     |  |     |
| Db | 186 | TTAGATTTATATATATAGAAAGAACATGTGGTCTTCTTATTCGATCGATGATTATTT      | 127 |
| QY | 681 | agctggtcccgcatgctgcatgcatcgttgcgtcgagacgagacaatcagacat         | 740 |
|    |     |  |     |
| Db | 126 | GCTCATATTAATTAATTCACAGATCCATCAATTTTCTTAAAGTTTGATTAATTTTCATTAT  | 67  |
| QY | 741 | caccacagcttctgcttctggagacaaataatgcttcttgaacaacaataataactata    | 800 |
|    |     |  |     |
| Db | 66  | GGCTTAATTAATTAATCTGTGTATGTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA   | 7   |

|    |     |    |     |
|----|-----|----|-----|
| QY | 801 | aa | 802 |
|    |     |    |     |
| Db | 6   | AA | 5   |

RESULT 4  
 US-08-361-441B-14/C  
 Sequence 14, Application US/08361441B  
 Patent No. 6077948  
 GENERAL INFORMATION:  
 APPLICANT: Russell, Mary E.  
 APPLICANT: Utans, Urike  
 TITLE OF INVENTION: MEDIATORS OF CHRONIC ALLOGRAFT REJECTION  
 NUMBER OF SEQUENCES: 43  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson P.C.  
 STREET: 225 Franklin Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110-2804  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows 95  
 SOFTWARE: FASTSEQ for Windows Version 2.0b  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/361,441B  
 FILING DATE: 21-DEC-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/171,385  
 FILING DATE: 21-DEC-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fraser, Janis K.  
 REGISTRATION NUMBER: 34,819  
 REFERENCE/DOCKET NUMBER: 05433/014001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617/542-5070

```

: TELEFAX: - 617/542-8906
:
: TELEEX: 200154
:
: INFORMATION FOR SEQ ID NO: 14:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 350 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: double
:
: TOPOLOGY: linear
:
US-08-361-441B-14

```

|                       |                |                |          |            |
|-----------------------|----------------|----------------|----------|------------|
| Query Match           | 3.48           | Score 34       | DB 3     | Length 350 |
| Best Local Similarity | 46.38          | Pred. No       | 0.27     |            |
| Matches 112           | Conservative 0 | Mismatches 130 | Indels 0 | Gaps 0     |

Qy 561 acaagaagatgcatgcatcagcttccattccatccatcgtgacaataatgaatgac 620  
 Db 246 AACTTATACCCACAGATTTTACATAGCCATATACATATATGTTTACCTAAACATTAAAG 187  
 Qy 621 ctacaatttcttcttggagcagaagtggtgtgctcttcttcttcaacgaaatccat 680  
 Db 186 TTAGATTTTATATATTAAGAAAGAACATGTGGTGTCTTCTTATCTGACATCGATPTATT 127  
 Qy 681 agctgcccgcacgctgcgacatctcgatgatcggtcggagacgacgacacagacac 740  
 Db 126 GCTCATATATAATATATTCACATCCATCCATCAATTTTCCCTAATAGTTTGATATTTCAATT 67  
 Qy 741 caccaacgcttctgtctgcggacacaataatgcttcttctgaaacaataataactata 800  
 Db 66 GGCTTAATTAATAATCTGTGTGTATGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 7

|    |     |    |     |
|----|-----|----|-----|
| QY | 801 | aa | 802 |
|    |     | 11 |     |
| Db | 6   | AA | 5   |

RESULT 5  
 US-08-458-434A-6/c  
 Sequence 6, Application US/08458434A  
 Patent No. 6083690  
 GENERAL INFORMATION:  
 APPLICANT: Harris Ph.D., Stephen E.  
 APPLICANT: Mundy M.D., Gregory R.  
 APPLICANT: Gosh-Choudhury Ph.D., Nandini  
 APPLICANT: Feng Ph.D., Jian Q.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING  
 TITLE OF INVENTION: OSTEOGENIC AGENTS  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: James C. Weseman, Esq.  
 STREET: 401 B. Street, Suite 1700  
 CITY: San Diego  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92101  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/458,434A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Weseman, James C.  
 REGISTRATION NUMBER: 30,507  
 REFERENCE/DOCKET NUMBER: P00060U50  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 699-3604  
 TELEFAX: 619-236-1048  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 15144 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-458-434A-6

Query Match  
Best Local Similarity 3.3%; Score 32.6; DB 3; Length 15144;  
Matches 74; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 466 catgcacattccaaacatgcacccgagatgcattgcagagtggaacggaactgaaa 525  
DB 11759 CAAGCACTTGTCTCAGACACTTCATGATAGCTTTGCATCTTAATACACACACTGAAA 11700  
QY 526 atcaaatcccatgtgtcagatgcagaagagatgcacactgaaagtatcgcttaacgat 585  
DB 11699 AACAACTACAAATTCCTTTTGGAAATTTATATACAAACCCGATATTCTTCGAT 11640

QY 586 ttcatatcacatcatgtacaat 608  
DB 11639 TACATATAAATACATTAATACTAT 11617

## RESULT 6

US-08-979-608A-16  
Sequence 16, Application US/08979608A  
Patent No. 6355451

GENERAL INFORMATION:  
APPLICANT: Lees, Ann M.  
Lees, Robert S.

Law, Simon W.  
Arlona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,608A  
FILING DATE: 26-NOV-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/048,547  
FILING DATE: 03-JUN-1997  
APPLICATION NUMBER: US 60/031,930  
FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1208 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: Coding Sequence

LOCATION: 1...651  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-08-979-608A-16

Query Match  
Best Local Similarity 3.2%; Score 32.2; DB 4; Length 1208;  
Matches 49; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 197 gaaagagctcacaagcagagagccacgacacagcttgagccctgcagacagcgagctgagtg 256  
DB 93 GAATGACCGCTCCGACAGGTGCGCAGCACCACGACTTAACGGAGCGGAGCTCAGAG 152  
QY 257 cgtgacacatggggtca 273  
DB 153 TGCCCAAGGAGAGGGTCA 169

## RESULT 7

US-08-814-095-7  
Sequence 7, Application US/08814095  
Patent No. 6025183

GENERAL INFORMATION:  
APPLICANT: Soreq, Hermona  
APPLICANT: Zakut, Haim  
APPLICANT: Shani, Moshe

TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KOHN & ASSOCIATES  
STREET: 30500 No. 6025183thwestern Highway, Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/814,095  
FILING DATE:

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Montgomery, Ilene N.  
REGISTRATION NUMBER: 38,972  
REFERENCE/DOCKET NUMBER: 2391.00066  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (248) 539-5050  
TELEFAX: (248) 539-5055  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Cosmid including Ache  
DESCRIPTION: promoter, Ache gene and Ars gene"  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 7q22

FEATURE:  
NAME/KEY: promoter  
LOCATION: 4089..22464  
OTHER INFORMATION: /function= "Ache Promotor"  
OTHER INFORMATION: /standard\_name= "Ache Promotor"

FEATURE:  
NAME/KEY: Coding Sequence

|    |                        |                                  |  |
|----|------------------------|----------------------------------|--|
| 1  | NAME/KEY:              | exon                             |  |
| 2  | LOCATION:              | 22465..22537                     |  |
| 3  | OTHER INFORMATION:     | /function= "non-translated"      |  |
| 4  | OTHER INFORMATION:     | /gene= "ACHE"                    |  |
| 5  | OTHER INFORMATION:     | /number= 1                       |  |
| 6  | FEATURE:               |                                  |  |
| 7  | NAME/KEY:              | exon                             |  |
| 8  | LOCATION:              | 24090..25177                     |  |
| 9  | IDENTIFICATION METHOD: | experimental                     |  |
| 10 | OTHER INFORMATION:     | /function= "(translation start:" |  |
| 11 | OTHER INFORMATION:     | /function= "4110)"               |  |
| 12 | OTHER INFORMATION:     | /evidence= EXPERIMENTAL          |  |
| 13 | OTHER INFORMATION:     | /gene= "ACHE"                    |  |
| 14 | OTHER INFORMATION:     | /number= 2                       |  |
| 15 | FEATURE:               |                                  |  |
| 16 | NAME/KEY:              | exon                             |  |
| 17 | LOCATION:              | 25524..26009                     |  |
| 18 | IDENTIFICATION METHOD: | experimental                     |  |
| 19 | OTHER INFORMATION:     | /evidence= EXPERIMENTAL          |  |
| 20 | OTHER INFORMATION:     | /gene= "ACHE"                    |  |
| 21 | OTHER INFORMATION:     | /number= 3                       |  |
| 22 | FEATURE:               |                                  |  |
| 23 | NAME/KEY:              | exon                             |  |
| 24 | LOCATION:              | 27005..27274                     |  |
| 25 | IDENTIFICATION METHOD: | experimental                     |  |
| 26 | OTHER INFORMATION:     | /evidence= EXPERIMENTAL          |  |
| 27 | OTHER INFORMATION:     | /gene= "ACHE"                    |  |
| 28 | OTHER INFORMATION:     | /number= 4                       |  |
| 29 | FEATURE:               |                                  |  |
| 30 | NAME/KEY:              | exon                             |  |
| 31 | LOCATION:              | 27255..28007                     |  |
| 32 | IDENTIFICATION METHOD: | experimental                     |  |
| 33 | OTHER INFORMATION:     | /evidence= EXPERIMENTAL          |  |
| 34 | OTHER INFORMATION:     | /gene= "ACHE"                    |  |
| 35 | OTHER INFORMATION:     | /number= 5                       |  |
| 36 | FEATURE:               |                                  |  |
| 37 | NAME/KEY:              | terminator                       |  |
| 38 | LOCATION:              | 27385..27387                     |  |
| 39 | FEATURE:               |                                  |  |
| 40 | NAME/KEY:              | exon                             |  |
| 41 | LOCATION:              | 28008..28129                     |  |
| 42 | IDENTIFICATION METHOD: | experimental                     |  |
| 43 | OTHER INFORMATION:     | /evidence= EXPERIMENTAL          |  |
| 44 | OTHER INFORMATION:     | /gene= "ACHE"                    |  |
| 45 | OTHER INFORMATION:     | /number= 6                       |  |
| 46 | FEATURE:               |                                  |  |
| 47 | NAME/KEY:              | terminator                       |  |
| 48 | LOCATION:              | 28129..28131                     |  |
| 49 | FEATURE:               |                                  |  |
| 50 | NAME/KEY:              | exon                             |  |
| 51 | LOCATION:              | complement (34528..34895)        |  |
| 52 | OTHER INFORMATION:     | /function= "arsenite resistance" |  |
| 53 | OTHER INFORMATION:     | gene= "gene"                     |  |
| 54 | OTHER INFORMATION:     | /gene= "AR"                      |  |
| 55 | OTHER INFORMATION:     | /number= 1                       |  |
| 56 | FEATURE:               |                                  |  |
| 57 | NAME/KEY:              | exon                             |  |
| 58 | LOCATION:              | complement (34092..34356)        |  |
| 59 | OTHER INFORMATION:     | /gene= "AR"                      |  |
| 60 | OTHER INFORMATION:     | /number= 2                       |  |
| 61 | FEATURE:               |                                  |  |
| 62 | NAME/KEY:              | exon                             |  |
| 63 | LOCATION:              | complement (33779..33963)        |  |
| 64 | OTHER INFORMATION:     | /gene= "AR"                      |  |
| 65 | OTHER INFORMATION:     | /number= 3                       |  |
| 66 | FEATURE:               |                                  |  |
| 67 | NAME/KEY:              | exon                             |  |
| 68 | LOCATION:              | complement (33493..33591)        |  |
| 69 | OTHER INFORMATION:     | /gene= "AR"                      |  |
| 70 | OTHER INFORMATION:     | /number= 4                       |  |
| 71 | FEATURE:               |                                  |  |
| 72 | NAME/KEY:              | exon                             |  |
| 73 | LOCATION:              | complement (33297..33408)        |  |

| Query Match   | Best Local Similarity | 3.2% | Score 31.8; | DB 3; | Length 35060; |
|---|-----------------------|------|-------------|-------|---------------|
| Matches 72; Conservative 0; Mismatches 67; Indels 0; Gaps 0;              |                       |      |             |       |               |
| QY 841 acgcgctccacagccgltgltgctgcatcgtccctccgcacgcagcgctgcctccacgg 900    |                       |      |             |       |               |
| DB 21250 AGGCGCTGCGGGGCGCTGCGGGGCGCGCGGCTTAGAGCGAGCGTGMGCGACGCGCC 21309   |                       |      |             |       |               |
| QY 901 tctgcctgctgctgcacactctgctgctgctgcgcgcagaaagggaagaaacagcgccg 960    |                       |      |             |       |               |
| DB 21310 ACCCCACGCGCGGCGCCGACCGCGCGCGCGGCGGACGACGCGCCGACGCTGGCTGGCG 21369 |                       |      |             |       |               |
| QY 961 cacacacactcacacacag 979  |                       |      |             |       |               |



```

: GENERAL INFORMATION:
: APPLICANT: KUBOTA, Michio
: APPLICANT: TSUSAKI, Keiji
: APPLICANT: HATTORI, Kazuko
: APPLICANT: SUGIMOTO, Toshiyuki
: TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 08/607,321
: FILING DATE: 26-FEB-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/399,646
: FILING DATE: 07-MAR-1995
: APPLICATION NUMBER: JP 59834
: FILING DATE: 07-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 59840
: FILING DATE: 07-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: BROWDY, Roger L.
: REGISTRATION NUMBER: 25,618
: REFERENCE/DOCKET NUMBER: KUBOTA-5
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: TELEX: 248633
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1767 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1767
: US-08-607-321-1

Query Match      3.1%; Score 31.2; DB 1; Length 1767;
Best Local Similarity 54.3%; Pred. No. 5.8;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 820 gctaacggcatgcccgaagcgcgtccacgcgtgtgttgcgactctgtctcccg 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1357 GCCACGGCCGAGGCGAGATGAGGAGTTCGAGCGCATGGGGTGGATCCCGCGTGTG 1416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 880 caccgacgctgcctccacgcgtcgcgtcgtcgcgcacacctgtgtgcgcgcg 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1417 CCCGATCCGACGATCCGAGACCTTCACCCGCTCCAAACTGAGTGGCGGAAGC 1472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-08-961-240-1
: Sequence 1, Application US/08961240
: Patent No. 5830715
: GENERAL INFORMATION:
: APPLICANT: KUBOTA, Michio
: APPLICANT: TSUSAKI, Keiji
: APPLICANT: HATTORI, Kazuko
: APPLICANT: SUGIMOTO, Toshiyuki

```

```

: APPLICANT: HATTORI, Kazuko
: APPLICANT: SUGIMOTO, Toshiyuki
: TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,240
: FILING DATE: 30-OCT-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/399,646
: FILING DATE: 07-MAR-1995
: APPLICATION NUMBER: JP 59834
: FILING DATE: 07-MAR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 59840
: FILING DATE: 07-MAR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: BROWDY, Roger L.
: REGISTRATION NUMBER: 25,618
: REFERENCE/DOCKET NUMBER: KUBOTA-5
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: TELEX: 248633
: INFORMATION FOR SRO ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1767 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1767
: US-08-961-240-1

Query Match      3.1%; Score 31.2; DB 2; Length 1767;
Best Local Similarity 54.3%; Pred. No. 5.8;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 820 gctaacggcatgcccgaagcgcgtccacgcgtgtgttgcgactctgtctcccg 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1357 GCCACGGCCGAGGCGAGATGAGGAGTTCGAGCGCATGGGGTGGATCCCGCGTGTG 1416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 880 caccgacgctgcctccacgcgtcgcgtcgtcgcgcacacctgtgtgcgcgcg 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1417 CCCGATCCGACGATCCGAGACCTTCACCCGCTCCAAACTGAGTGGCGGAAGC 1472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-08-605-501-1
: Sequence 1, Application US/08605501
: Patent No. 5834287
: GENERAL INFORMATION:
: APPLICANT: KUBOTA, Michio
: APPLICANT: TSUSAKI, Keiji
: APPLICANT: HATTORI, Kazuko
: APPLICANT: SUGIMOTO, Toshiyuki
: TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND

```





```

;
;   ZIP: 20004
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/607,321
;   FILING DATE: 26-FEB-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/399,646
;   FILING DATE: 07-MAR-1995
;   APPLICATION NUMBER: JP 59834
;   FILING DATE: 07-MAR-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 59840
;   FILING DATE: 07-MAR-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: BROWDY, Roger L.
;   REGISTRATION NUMBER: 25,618
;   REFERENCE/DOCKET NUMBER: KUBOTA-5
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-628-5197
;   TELEFAX: 202-737-3528
;   TELEX: 248633
;   INFORMATION FOR SEQ. ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2161 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 207..1994
;
US-08-607-321-11

Query Match          3.1%; Score 31.2; DB 1; Length 2161;
Best Local Similarity 54.3%; Pred. No. 6.6;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 820 gctaacgcatgacgagtaaacgcgtccacgcgtgtgttgatctcgtctccg 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1584 GCCACGCCGACGAGGACGAGATTCGACGCGATGGGTGGATCCCGCGTGTG 1643

QY 880 cagcagcgctgcctccacgcgtcgtcgtccacactcgtgtgcgcgcg 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1644 CCCGATCCGACGATCCGAGACCTTCACCCGCTCCAACTGACTGGCGGGAAGC 1699

RESULT 15
US-08-961-240-11
; Sequence 11, Application US/08961240
; Patent No. 5830715
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

;
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/961,240
;   FILING DATE: 30-OCT-1997
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/399,646
;   FILING DATE: 07-MAR-1995
;   APPLICATION NUMBER: JP 59834
;   FILING DATE: 07-MAR-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 59840
;   FILING DATE: 07-MAR-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: BROWDY, Roger L.
;   REGISTRATION NUMBER: 25,618
;   REFERENCE/DOCKET NUMBER: KUBOTA-5
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-628-5197
;   TELEFAX: 202-737-3528
;   TELEX: 248633
;   INFORMATION FOR SEQ. ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2161 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 207..1994
;
US-08-961-240-11

```

```

Query Match          3.1%; Score 31.2; DB 2; Length 2161;
Best Local Similarity 54.3%; Pred. No. 6.6;
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 820 gctaacgcatgacgagtaaacgcgtccacgcgtgtgttgatctcgtctccg 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1584 GCCACGCCGACGAGGACGAGATTCGACGCGATGGGTGGATCCCGCGTGTG 1643

QY 880 cagcagcgctgcctccacgcgtcgtcgtccacactcgtgtgcgcgcg 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1644 CCCGATCCGACGATCCGAGACCTTCACCCGCTCCAACTGACTGGCGGGAAGC 1699

```

Search completed: September 12, 2002, 08:12:27  
Job time: 14804 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 03:32:50 : Search time 3207.55 Seconds  
(without alignments)  
4207.871 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_1\_1000  
Perfect score: 1000  
Sequence: 1 agaacacctcatataga.....acatccccgtggtccct 1000

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vtl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 47.6  | 4.8         | 1101   | 12    | CNS00LT2    |
| 2          | 47.7  | 4.7         | 1101   | 9     | AL513871    |
| 3          | 43.4  | 4.3         | 844    | 12    | CNS0052P    |
| 4          | 40.4  | 4.0         | 931    | 10    | BF142642    |
| 5          | 38.6  | 3.9         | 427    | 10    | BG895352    |
| 6          | 38.6  | 3.9         | 1031   | 10    | W29451      |
| 7          | 38.4  | 3.8         | 866    | 10    | BI953458    |
| 8          | 38    | 3.8         | 541    | 12    | BH581366    |
| 9          | 38    | 3.8         | 1098   | 10    | BF980551    |
| 10         | 37.8  | 3.8         | 1101   | 12    | CNS0039G    |
| 11         | 37.8  | 3.8         | 725    | 12    | AQ421725    |
| 12         | 37.8  | 3.8         | 945    | 9     | AL544411    |
| 13         | 37.8  | 3.8         | 1281   | 12    | AG075401    |
| 14         | 37.6  | 3.8         | 1477   | 12    | AG043219    |
| 15         | 37.4  | 3.7         | 471    | 12    | AZ113002    |
| 16         | 37.4  | 3.7         | 1101   | 12    | CNS017V6    |
| 17         | 37.2  | 3.7         | 417    | 10    | H46465      |

|    |      |     |      |    |          |
|----|------|-----|------|----|----------|
| 18 | 37.2 | 3.7 | 509  | 12 | AQ123336 |
| 19 | 37.2 | 3.7 | 608  | 12 | BH113363 |
| 20 | 37   | 3.7 | 238  | 9  | BE183329 |
| 21 | 37   | 3.7 | 253  | 12 | BH292353 |
| 22 | 37   | 3.7 | 350  | 12 | CNS02E4E |
| 23 | 37   | 3.7 | 555  | 9  | AV441208 |
| 24 | 37   | 3.7 | 769  | 12 | CNS01WN  |
| 25 | 37   | 3.7 | 775  | 10 | BE909760 |
| 26 | 37   | 3.7 | 1006 | 12 | CNS00JWJ |
| 27 | 36.8 | 3.7 | 273  | 9  | BB514731 |
| 28 | 36.8 | 3.7 | 441  | 9  | AI030202 |
| 29 | 36.6 | 3.7 | 696  | 10 | BI154298 |
| 30 | 36.6 | 3.7 | 790  | 10 | BF182476 |
| 31 | 36.6 | 3.7 | 839  | 12 | CNS004NB |
| 32 | 36.6 | 3.7 | 908  | 12 | AZ549202 |
| 33 | 36.6 | 3.7 | 1101 | 12 | CNS00DY8 |
| 34 | 36.6 | 3.7 | 1101 | 12 | CNS0167A |
| 35 | 36.4 | 3.6 | 182  | 9  | AI611686 |
| 36 | 36.4 | 3.6 | 303  | 12 | CNS00612 |
| 37 | 36.4 | 3.6 | 519  | 10 | C98216   |
| 38 | 36.4 | 3.6 | 737  | 10 | BF294776 |
| 39 | 36.4 | 3.6 | 861  | 9  | AL548346 |
| 40 | 36.4 | 3.6 | 997  | 12 | CNS006DN |
| 41 | 36.4 | 3.6 | 1010 | 12 | CNS0167E |
| 42 | 36.4 | 3.6 | 1022 | 12 | CNS003RM |
| 43 | 36.2 | 3.6 | 438  | 10 | BM090853 |
| 44 | 36.2 | 3.6 | 855  | 10 | BI099897 |
| 45 | 36.2 | 3.6 | 1594 | 10 | BE962092 |

#### ALIGNMENTS

RESULT 1  
CNS00LT2/c  
LOCUS  
DEFINITION  
CNS00LT2 1101 bp DNA linear GSS 14-JUN-1999  
Drosophila melanogaster genome survey sequence TET3 end of BAC:  
BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.

ACCESSION  
AL078714.1 GI:5102004  
VERSION  
AL078714  
KEYWORDS  
GSS.  
SOURCE  
fruit fly.  
ORGANISM  
Drosophila melanogaster

REFERENCE  
AUTHORS  
JOURNAL  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

#### COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

#### FEATURES

Location/Qualifiers  
1..1101  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone\_lib="RPCI-98"

| BASE COUNT | ORIGIN                    |
|------------|---------------------------|
| 469 a      | 6 c 69 g 151 t 406 others |

|                       |                  |                |          |             |
|-----------------------|------------------|----------------|----------|-------------|
| Query Match           | 4.8%             | Score 47.6     | DB 12    | Length 1101 |
| Best Local Similarity | 23.6%            | Pred. No. 0.43 |          |             |
| Matches 131           | Conservative 160 | Mismatches 261 | Indels 4 | Gaps 1      |

**Qy** 110 aaagcgctcaaccagtcacaaacytctcyctagaatcaccaagctgcgaagttaagcgcgag 169  
::: : :::: | :|: | :| : : : : : : : :  
**Db** 999 MAAMNMNMNMMAATTAHNTTTTHPTTYMAMCTTHTMMNMNMMAAMAMNMMSMC 9400

**Dy** 170 accaccacaaacagcgcattgcaactctgacagatgtccacgcagaagcccacccaag 229  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
**Db** 939 MCMNCMNMNCCCCMCMNMNMNMNTTTTNNHMCMCMCMCCMCMCMNMNMNMKAMMM 880

Dy 230 gcttgagcctgcacacggaactgtatgctgatcacatgggtcatcatatggcgctcgag 289  
|||::|:::|:::|:::|:::|:::|:::  
Db 879 mmtttttnnccmmmmmmmmmaammammmnmcmhmamttttttttnnammnahtnttm 820

QY 250 caatgaagagaacgcacatacaacacatgatgtctaacgacctgatgaggagca 349  
:: :: : | | :: :: : | : | :  
Db 819 MMTTMMMDNMCMNMAAATTTMNNMMAMHNTTNHTMNCMCNCCMNMMNNNNCSM 760

QY 350 acctgtcaccttttccctcgtgaatcatagctcacacttttttaatgaaagaag 409  
::: |:: :::: | :::: | :::: | :::: | ::::  
Db 759 mcmcmcmcmcmcmymnnnhttwttmnmnmamattmttcmmntnkmnnhtttmmmmma 700

Qy 410 agtgcgcaaacacatgcattctcaacagaataaattctcaaccaccatgcagt 469  
| :: | : | | | | | | | : | : | : |  
Db 699 AAWWTTTWAAHAAWTTTTTATTTAAATAAAAAAMMTAAAAAATWTANAAG 6400

Oy 470 caattctcaaacatgcacccgagcagtcgatcgaggttgyaaacgaagaactgaaatca 529  
|| ||| : | | | : : | : | : | : |  
Db 639 AAAAATTWAAAATAAAWAAWMAAATGHHWTWWNV---TWAMMTAAWMAAAAWATTAA 584

OY 530 acatcccaattgtcgaagtcgaaagagatgcacactgaaagtatcgattacgatttca 589  
| | | : | : | : | | | : |  
Db 583 AAAAATTTTATTAMTANTMMANHAMATGMAAAAMMAAAAAATATATCTAATWMAAWMTAAA 524

OY 590 ttcacatacagtaacaatacataatgaccaccaattgttttggagcagatggt 649  
:: ::::| : || | :::: | || | ||||  
Db 523 ahwtttwmnhhtthaalawaalaattmmwaaaaawaatTTTTNTTTTNTTTT 464

QY 650 gtggtcttttttttc 665  
| | | | | | | | | |  
Db 463 TTTTTTTTTTTTTTTT 448

|          |   |
|----------|---|
| RESULT   | 2 |
| AL513871 |   |

|            |  |         |      |        |               |
|------------|--|---------|------|--------|---------------|
| LOCUS      | AL513871   | 1101 bp | mRNA | linear | EST 13-FEB-96 |
| DEFINITION | AL513871 LTI_NFL006_PL2 Homo sapiens CDNA clone CL0BA002ZA07 3 prime, mRNA sequence. |         |      |        |               |
| ACCESSION  | AL513871   |         |      |        |               |

|          |              |             |
|----------|--------------|-------------|
| VERSION  | AL5J138/1.1  | GI:12111363 |
| KEYWORDS | EST.         |             |
| SOURCE   | human.       |             |
| ORGANISM | Homo sapiens |             |

REFERENCE  
AUTHORS  
Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.  
1 (bases 1 to 1101)  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi  
1 (bases 1 to 1101)

|         |  |
|---------|--|
| TITLE   | Full-length cDNA libraries and normalization |
| JOURNAL | Unpublished (2001)                           |
| COMMENT | Genoscope - Centre National de Sequencage    |

BP 191 91006 EVRY cedex - France  
Email: [segrete@genoscope.cns.fr](mailto:segrete@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).  
Location/Qualifiers  
1 1101

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="Cl0BA0022A07"  
/clone_1lb="LT1_NFL006_PL2"
```

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6

Life Technologies. Contact : Feng Liang Life Technologies,  
a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

http://fulllength.invitrogen.com"  
a 33 c 248 g 323 t 335 others

4.7%; Score 47; DB 9; length 1101;  
24.1%; Pred. No. 0.61;

ttttttgtctcttctcgacggtggtcgtggaagattagcgtctagt 73

aacagccatttagccctgcttacaagaagctaacccagtcacaacg 133  
 ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: |  
 :::::

ccaccagctgcgaagttaagcgcgagacacccaaacagcgcattcgaa 193  
::: |: | : :: | | ::::::::::::::: | ::: :::

ctcaccacgacgagacccacacacacgacctgacacgacgacgtga 253  
:: ||| :::: ||| :::: :: : :::: ||

atgggg 270  
||||  
|||

844 bp DNA linear GSS 03-JUN-1999

1 GI:4932342

a melanogaster  
 a, Melazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 a, Neoptera; Endopterygota; Diptera; Brachycera;

submission  
2.  
31 to 844)  
Ephedroidea; Drosophilidae; Drosophila.

tion of this Rac-end sequence was carried out as part of a  
 (02-JUN-1999) Genoscope - Centre National de Séquençage :  
 0006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 www.genoscope.cns.fr)

tion with the Berkeley Drosophila Genome Project (BDGP). We are constructing a physical map of the *Drosophila* ester genome using these BACs. For further information [http://www.fruitfly.org/The BDGP](http://www.fruitfly.org/The%20BDGP)

After BAC library was prepared by Kazutoyo Oseegawa and Pieter de Jong's laboratory in the department of chemistry at the Roswell Park Cancer Institute in Buffalo, New York, we cloned pD1-08 and used constructed by partial







```

BASE COUNT      270 a      424 c      220 g      184 t
ORIGIN
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-gt primed. Average insert size 1,767 kb. Library
enriched for full-length clones and constructed by Life
Technologies." Note: this is a NIH-MGC Library."

```

|                       |       |              |    |            |    |        |      |
|-----------------------|-------|--------------|----|------------|----|--------|------|
| Query Match           | 3.8%  | Score        | 38 | DB         | 10 | Length | 1098 |
| Best Local Similarity | 60.8% | Pred. No.    | 96 |            |    |        |      |
| Matches               | 62    | Conservative | 0  | Mismatches | 40 | Indels | 0    |
|                       |       |              |    |            |    | Gaps   | 0    |

**QY** 888 gtgagctccacacgtcgcgtcgctgcgtgcacactctgctgtgcgcgcgaagaagga 947  
+  
**Db** 902 GACTCCCGCCCCGACCTCCCCCGGTCTCCCAACCCTTCGCCCCACACCAACGGCGC 9611

|    |   |
|----|---|
| Qy | 948 agaacgaacgcgcacacacactcacacagcgcaactcccc 989<br>   <br> |
| Db | 962 ACACCGCCGCGCACCCACCACGCCACGCACACACCCCC 1003   |

| RESULT     | 10   |
|------------|--|
| CNS00396   |  |
| LOCUS      |  |
| DEFINITION |  |
| CNS00396   | 1101 bp DNA linear GSS 03-JUN-1999                               |
|            | Drosophila melanogaster genome survey sequence TE13 end of BAC # |
|            | BACR08K10 of RP11-98 library from Drosophila melanogaster (fruit |
|            | fly), genomic survey sequence.                                   |
|            | 11027031   |
| ACCESSION  |  |

|           |                                     |            |
|-----------|-------------------------------------|------------|
| ACCESSION | AL063921                            |            |
| VERSION   | AL063921.1                          | GI:4941778 |
| KEYWORDS  | GSS.                                |            |
| COMMENT   | GenBank accession number AL063921.1 |            |

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequen

JOURNAL  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a

melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila\\_melanogaster\\_BAC\\_library](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library) was prepared by Kazutoyo Osoegawa and Aaron Mammoso in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source   | 1. . 1101           |

|            |       |      |       |       |            |
|------------|-------|------|-------|-------|------------|
| BASE COUNT | 201 a | 64 c | 131 g | 202 t | 503 others |
| ORIGIN     |       |      |       |       |            |

|                       |       |              |     |            |      |        |      |
|-----------------------|-------|--------------|-----|------------|------|--------|------|
| Query Match           | 3.8%  | Score        | 38  | DB         | 12   | Length | 1101 |
| Best Local Similarity | 14.0% | Pred. No.    | 96  |            |      |        |      |
| Matches               | 37    | Conservative | 132 | Mismatches | 95   | Indels | 0    |
|                       |       |              |     |            | Gaps |        | 0    |

QY 359 ctttccctctggaatcatagctcacctttttttaatgtgaagcaagattgcaa 418

[illegible]

Db 1073 WWAWMAMHNNMUAHYHWAHHCW 1096

|            |  |
|------------|--|
| RESULT     | 11   |
| AQ421725   |  |
| LOCUS      |  |
| DEFINITION | AQ421725 725 bp DNA linear GSS 23-MAR-1999<br>RCPI-11-190G1.TV RCPI-11 Homo sapiens genomic clone RCPI-11-190G1, |
| ACCESSION  | DNA sequence.  |
| VERSION    | AQ421725   |
| KEYWORDS   | AQ421725.1 GI:4479449  |
| SOURCE     | GSS.<br>human.   |

| ORGANISM  | REFERENCE          |
|---|--------------------|
| <i>Homo sapiens</i>   | 1 (bases 1 to 725) |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |                    |
| Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.         |                    |

**AUTHORS** Zhao, S., Adams, M. D., Niemerman, W., Malek, J., de Jong, P. and Venter, J. C.  
**TITLE** Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building  
**JOURNAL** Unpublished (1997)  
**COMMENT** Contact: Shaying Zhao, William Niemerman, Mark Adams

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel.: 301 838 0200  
Fax: 301 838 0208

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong ([pieterdejong.med.buffalo.edu](mailto:pieterdejong.med.buffalo.edu)). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs ([info@resgen.com](mailto:info@resgen.com)). BAC end search page: [http://www.tigr.org/cdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html)  
Seq primer: 17  
Class: BAC ends.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source   | 1. .725             |

```

/organism="Homo sapiens"
/db xref="GDB:7572720"
/db xref="taxon:9606"
/clone="RPCT-11-19031"
/clone_1fb="RPCT-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCT11 Human Male BAC Library"
BASE COUNT      256 a      123 c      129 g      216 t      1 others
ORIGIN

```

|                           |       |                   |          |            |
|---------------------------|-------|-------------------|----------|------------|
| Query Match               | 3.8%  | Score 37.8        | DB 12    | Length 725 |
| Best Local Similarity     | 48.0% | Pred. No. 1.1e+02 |          |            |
| Matches 108; Conservative | 0     | Mismatches 117    | Indels 0 | Gaps 0     |

QY 449 ttctcaaacccatgacatgcaatttctcaaacatgcaaccgagagtcctcatgcgagtg 508





|  |  |  |  |  |  |
|--|--|--|--|--|--|
| ORGANISM   | BAC library clone: PTB-021H12.F.   |  |  |  |  |
| REFERENCE  | Pan troglodytes  |  |  |  |  |
| AUTHORS  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.   |  |  |  |  |
| TITLE  | 1 (sites)  |  |  |  |  |
| REFERENCE  | Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.   |  |  |  |  |
| TITLE  | 2 (bases 1 to 1477)  |  |  |  |  |
| REFERENCE  | Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.   |  |  |  |  |
| TITLE  | Direct Submission  |  |  |  |  |
| JOURNAL  | Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Saito-cho,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@sc.riken.go.jp, URL:http://ngp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170) |  |  |  |  |
| COMMENT  | Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.  |  |  |  |  |
| PRIMERS  | Sequencing: -21M13   |  |  |  |  |
| LIBRARY  | Vector : pKS145  |  |  |  |  |
| FEATURES   | R.Site 1 : SacI  |  |  |  |  |
| source   | R.Site 2 : SacI.   |  |  |  |  |
| location/Qualifiers  | 1..1477  |  |  |  |  |
| organism="Pan troglodytes"   | /db_xref="taxon:9598"  |  |  |  |  |
| clone="PTB-021H12.F"   | /sex="male"  |  |  |  |  |
| cell_type="lymphoblast"  | /clone_lib="PTB Chimpanzee Male BAC Library"   |  |  |  |  |
| BASE COUNT   | 446 a 382 c 525 g 83 t 41 others   |  |  |  |  |
| ORIGIN   |  |  |  |  |  |
| Query Match  | 3.8%; Score 37.6; DB 12; Length 1477;  |  |  |  |  |
| Best Local Similarity  | 55.1%; Pred.No.1.2e+02;  |  |  |  |  |
| Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;       |  |  |  |  |  |
| QY 808 gtactagagggccgcgtcaagggcatggcgaagagcgtccacgcgtgttgcat 867   |  |  |  |  |  |
| Db 937 gTGCCTGGTGGCCCGGGGCGTGGCGGCGTGTGCGCTTCCGGCGCTTNTTTCGCCG 878 |  |  |  |  |  |
| QY 868 ctctgcctccgcacagcagcgtcgcctcacccgctcgtcgcgtgcacgtcgtcgt 927 |  |  |  |  |  |
| Db 877 CGGCCCGCCCGCTCGCGCGCGCGCTTCCGCTGCGCTGCTGCTGCTGCTTTC 818     |  |  |  |  |  |
| QY 928 ggcgcgcg 934  |  |  |  |  |  |
| Db 817 GCGCGCG 811   |  |  |  |  |  |
| RESULT 15  |  |  |  |  |  |
| AZ113002/c   |  |  |  |  |  |
| LOCUS  | AZ113002 471 bp DNA linear GSS 12-MAY-2000   |  |  |  |  |
| DEFINITION   | RPci-23-13D17.TJ RPci-23 Mus musculus genomic clone RPci-23-13D17,   |  |  |  |  |
| ACCESSION  | AZ113002   |  |  |  |  |
| VERSION  | AZ113002.1 GI:7772220  |  |  |  |  |
| KEYWORDS   | GSS.   |  |  |  |  |
| SOURCE   | house mouse.   |  |  |  |  |
| ORGANISM   | Mus musculus   |  |  |  |  |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  |  |  |  |  |
| AUTHORS  | Zhao,S., Niemman,W., Feldblum,T., Malek,J., Shatsman,S., Akiret ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.  |  |  |  |  |
| TITLE  | Mouse BAC End Sequences from Library RPci-23   |  |  |  |  |

```
JOURNAL COMMENT Unpublished (1999) Other.GSSS: RPCI-23-13D17.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhaoe@tigr.org Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pjeter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac\_ends/mouse/BAC\_end\_intro.html Plate: 13 row: D column: 17 Seq primer: SP6 Class: BAC ends.
```

|                       | location/qualifiers  |
|-----------------------|--|
| SOURCE                | . .471<br>/organism="Mus musculus"<br>/strain="C57BL/6J"<br>/db_xref="taxon:10090"<br>/clone="RPCI-23-13D17"<br>/_clone_id="RPCI-23"<br>/_sex="Female"<br>/_lab_host="DH10B"<br>/_note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:<br>EcoRI_Site_2: EcoRI; Female C57Bl/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)." |
| BASE COUNT            | 132 a 129 c 97 g 113 t   |
| ORIGIN                |  |
| Query Match           | 3.7%; Score 37.4; DB 12; Length 471;   |
| Best Local Similarity | 62.1% ; Pred. No. 1.3e+02;   |
| Matches               | 59; Conservative 0; Mismatches 36; Indels 0; Gaps 0;   |
| OY                    | 375 attcatagctcacactttttttaatggaagaagcagggtggcaacaacatacgattcca 434<br>  |
| Dd                    | 467 ATTCTCGGTCAGCCCATTTTATTATGTAAACAATAAGITTAGATCCATGCCTTGTTGG 408   |
| OY                    | 435 acaaaggaaaattaatlctccaacaccacgatgcataly 469<br>  |
| Dd                    | 407 AGGAAGAAGTTAAGAACCAACCATCTCGGTTG 373   |

Search completed: September 12, 2002, 08:10:10  
Job time: 16640 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 07:27:09 ; Search time 3917.59 Seconds  
(without alignments)  
5347.028 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_4500\_5500  
1001  
Sequence: 1 tactagcttaagattccca.....tgtattaacacagttgcaca 1001

## Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Database :

Listing first 45 summaries

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query Score Match Length DB ID Description  
-----  
SUMMARIES

|    |       |       |        |   |            |                      |
|----|-------|-------|--------|---|------------|----------------------|
| 1  | 1001  | 100.0 | 11463  | 6 | AX031271   | AX031271 Sequence    |
| 2  | 965   | 96.4  | 11475  | 8 | AF338431   | AF338431 Aegilops    |
| 3  | 141.2 | 14.1  | 7020   | 8 | ATSB21     | AT000497 Arabidops   |
| 4  | 141.2 | 14.1  | 101365 | 8 | AC006919   | AC006919 Arabidops   |
| 5  | 131.6 | 13.1  | 102887 | 8 | ATP17C15   | AL162506 Arabidops   |
| 6  | 128.2 | 12.8  | 2853   | 8 | TAU66376   | U66376 Trillium ae   |
| 7  | 125.8 | 12.6  | 2554   | 8 | AF064560   | AF064560 Hordeum v   |
| 8  | 124.2 | 12.4  | 2726   | 8 | AF338432   | AF338432 Trillium    |
| 9  | 124.2 | 12.4  | 2970   | 8 | AF286319   | AF286319 Trillium    |
| 10 | 124.2 | 12.4  | 2970   | 8 | TASBA2     | Y11282 T. aestivum   |
| 11 | 111.6 | 11.1  | 23449  | 8 | AF072725   | AF072725 Zea mays    |
| 12 | 109   | 10.9  | 2364   | 6 | E14724     | E14724 Rice mRNA f   |
| 13 | 109   | 10.9  | 3015   | 6 | E14723     | E14723 Rice mRNA f   |
| 14 | 109   | 10.9  | 3015   | 8 | AB023498   | AB023498 Oryza sat   |
| 15 | 105.8 | 10.6  | 2795   | 8 | ZM059448   | ZM059448 Zea mays st |
| 16 | 105.2 | 10.5  | 2780   | 8 | AF064561   | AF064561 Hordeum v   |
| 17 | 100.8 | 10.1  | 2725   | 8 | AR106495   | AR106495 Sequence    |
| 18 | 100.8 | 10.1  | 2725   | 8 | M2EGUCRTRN | AR106495 Sequence    |
| 19 | 97.6  | 9.8   | 2918   | 8 | R1CBCE3    | L08065 Corn starch   |
| 20 | 97.6  | 9.8   | 2919   | 6 | E08183     | D16201 Rice mRNA f   |
| 21 | 91    | 9.1   | 3050   | 6 | A92164     | E08183 Gene of sta   |
| 22 | 89.6  | 9.0   | 2493   | 8 | STSEBIT    | A92164 Sequence 30   |
| 23 | 89.6  | 9.0   | 2563   | 6 | AX256072   | AX256072 Solanum t   |
| 24 | 89.6  | 9.0   | 3549   | 8 | PSSEBIGN   | AX256072 Sequence    |
| 25 | 89.4  | 8.9   | 2517   | 8 | AB042937   | X80009 P. sativum m  |
| 26 | 89.4  | 8.9   | 3123   | 8 | AB071286   | AB042937 Ipomoea b   |
| 27 | 88.6  | 8.8   | 2668   | 8 | ATU18817   | AB071286 Ipomoea b   |
| 28 | 88    | 8.8   | 2523   | 8 | STU011869  | U18817 Arabidops     |
| 29 | 88    | 8.8   | 2529   | 6 | A58167     | AJ011889 Solanum t   |
| 30 | 88    | 8.8   | 2576   | 6 | A58166     | A58167 Sequence 17   |
| 31 | 88    | 8.8   | 2578   | 6 | A58169     | A58166 Sequence 16   |
| 32 | 88    | 8.8   | 2962   | 8 | STU011888  | A58169 Sequence 19   |
| 33 | 88    | 8.8   | 2992   | 8 | STU011890  | AJ011888 Solanum t   |
| 34 | 88    | 8.8   | 3003   | 6 | A58162     | AJ011890 Solanum t   |
| 35 | 88    | 8.8   | 3033   | 6 | A58164     | A58162 Sequence 12   |
| 36 | 88    | 8.8   | 3231   | 6 | A58168     | A58164 Sequence 14   |
| 37 | 87.6  | 8.8   | 3074   | 6 | AR123355   | A58168 Sequence 18   |
| 38 | 86    | 8.6   | 1393   | 6 | AR123356   | AR123355 Sequence    |
| 39 | 85.2  | 8.5   | 3360   | 8 | AB029548   | AR123356 Sequence    |
| 40 | 84.8  | 8.5   | 2955   | 8 | STU011885  | AB029548 Phaseolus   |
| 41 | 84.8  | 8.5   | 2975   | 6 | A58163     | AJ011885 Solanum t   |
| 42 | 84.4  | 8.4   | 2542   | 8 | ATU22428   | A58163 Sequence 13   |
| 43 | 83    | 8.3   | 2913   | 6 | A92162     | U22428 Arabidops     |
| 44 | 80.4  | 8.0   | 1481   | 8 | STU011891  | A92162 Sequence 28   |
| 45 | 74.2  | 7.4   | 2843   | 8 | MESBERNA   | AJ011891 Solanum t   |

## ALIGNMENTS

RESULT 1  
AX031271 LOCUS AX031271 11463 bp DNA linear PAT 20-SEP-2000

DEFINITION Sequence 10 from Patent WO914314.  
ACCESSION AX031271  
VERSION AX031271.1 GI:10278603

KEYWORDS  
SOURCE  
ORGANISM

Aegilops tauschii.  
Aegilops tauschii.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 11463)  
Il, Z., Morell, M. and Rahman, S.  
Regulation of gene expression in plants  
Patent: WO 914314-A 10 25-MAR-1999;  
GOODMAN FIELDER LTD (AU) ; LI ZHONGYI (AU) ; MORELL MATTHEW (AU) ;  
RAHMAN SADEOUR (AU) ; UNIV AUSTRALIAN (AU) ; COMM SCI ENT IND RES  
ORG (AU) ; GROUPE LIMAGRAIN PACIFIC PTY L (AU)

FEATURES  
source  
1..11463  
/organism="Aegilops tauschii"  
/db\_xref="taxon:37682"





```

            /gene="Sbe2.1"
            /number=2
            /number=1751.1848
            /gene="Sbe2.1"
            /number=2
            /number=1849.2084
            /gene="Sbe2.1"
            /number=3
            /number=2085.2276
            /gene="Sbe2.1"
            /number=3
            /number=2277.2375
            /gene="Sbe2.1"
            /number=4
            /number=2376.2531
            /gene="Sbe2.1"
            /number=4
            /number=2532.2574
            /gene="Sbe2.1"
            /number=5
            /number=2575.3044
            /gene="Sbe2.1"
            /number=5
            /number=3045.3104
            /gene="Sbe2.1"
            /number=6
            /number=3105.3290
            /gene="Sbe2.1"
            /number=6
            /number=3291.3371
            /gene="Sbe2.1"
            /number=7
            /number=3372.3460
            /gene="Sbe2.1"
            /number=7
            /number=3461.3577
            /gene="Sbe2.1"
            /number=8
            /number=3578.3792
            /gene="Sbe2.1"
            /number=8
            /number=3793.3876
            /gene="Sbe2.1"
            /number=9
            /number=3877.4233
            /gene="Sbe2.1"
            /number=9
            /number=4234.4355
            /gene="Sbe2.1"
            /number=10
            /number=4356.4500
            /gene="Sbe2.1"
            /number=10
            /number=4501.4620
            /gene="Sbe2.1"
            /number=11
            /number=4621.4715
            /gene="Sbe2.1"
            /number=11
            /number=4716.4845
            /gene="Sbe2.1"
            /number=12
            /number=4846.4964
            /gene="Sbe2.1"
            /number=12
            /number=4965.5075
            /gene="Sbe2.1"
            /number=13
            /number=5076.5192
            /gene="Sbe2.1"
            /number=13
            /number=5193.5321
            /gene="Sbe2.1"

```

```

            /number=14
            /gene="Sbe2.1"
            /number=14
            /number=5404.5507
            /gene="Sbe2.1"
            /number=15
            /number=5508.5601
            /gene="Sbe2.1"
            /number=15
            /number=5602.5746
            /gene="Sbe2.1"
            /number=16
            /number=5747.5877
            /gene="Sbe2.1"
            /number=16
            /number=5878.6435
            /gene="Sbe2.1"
            /number=17
            /number=6436.6519
            /gene="Sbe2.1"
            /number=17
            /number=6520.6663
            /gene="Sbe2.1"
            /number=18
            /number=6781
            /gene="Sbe2.1"

BASE COUNT      1955 a      1249 c      1406 g      2410 t
ORIGIN

```

```

Query Match      14.1%; Score 141.2; DB 8; Length 7020;
Best Local Similarity 62.2%; Pred. No. 8.5e-22;
Matches 270; Conservative 0; Mismatches 138; Indels 26; Gaps 2;

```

```

QY      313 caggaaacgaaataatcattacataatgtaatttaaggaatgaggtgttgcacaaattaa 372
            |||||||
DB      4231 CAGGAACCAAAAGATAAATACATATGCCAATTAGAGATGATCTCCCGTATMAAA 4290
QY      373 aggccttgatacaatgcagtgagataatggcaatccaggaatcattacataatgcaagc 432
            |||||
DB      4291 AAGCTAGGCTAATATGCTGTGACAGATATGCGCATTCAGAGCAAGCCTACTATGCCAGC 4350
QY      433 ttgggtaltcacacaatccattttctctgatacaactctaccatltggagcatt 492
            |||||||
DB      4351 TTGGGTAT-----CTTCCTTCACAGATATTAACATGTCTACCTATCATATT 4400
QY      493 acatccataatgctatgcacataaataattggatataatccttattatagatatagt 552
            |||||
DB      4401 AATGAATACCTGTATAAAGCCATTGTGATTCTTCAGAGTC-----GT 4444
QY      553 acaactacactagattcttcgaaagaatcatttattgtgtgtgtgttcacagttac 612
            |||||
DB      4445 AAACCTATTTCCTCCCTTGATGAATGACAGATGCTATTGCTATATGCAATGACAGTAT 4504
QY      613 catgttactaatlttttgcaccaagtagccgttttggaaactccaggaacttaaatcc 672
            |||||
DB      4505 CATGTGACAAATTTTTCGACCTAGCCGCTTTGGAAACACGATGATGCTTAATCT 4564
QY      673 ttgatacgaatagacacagagcttggttctgtctgttattagatatgttcataatgtaa 732
            |||||
DB      4565 TTGATAGACAAAGCTCATGAGCTAGTGTGTTGTTCGATGATATGTGCACAGGTAT 4624
QY      733 ttagtccaattaa 746
            |||||
DB      4625 TCGCTTCTCTTTAA 4638

```

```

RESULT      4
AC006919      101365 bp      DNA      linear      PLN 05-APR-2000
LOCUS      AC006919
DEFINITION      Arabidopsis thaliana chromosome II section 199 of 255 of the
                  complete sequence. Sequence from clones F2H17, F1011, F13K3.

```



ACCESSION AC006919 AE002093  
VERSION AC006919.5 GI:6598632  
KEYWORDS HTG.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 101365)  
Lin.X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,  
Fuji,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,  
Felibuyum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Ronning,C.M.,  
Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., Vanaken,S.E., Umayam,L.,  
Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,  
Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D.,  
Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and  
Venter,J.C.  
TITLE Sequence and analysis of chromosome 2 of the plant Arabidopsis.  
JOURNAL thaliana  
MEDLINE Nature 402 (6763), 761-768 (1999)  
PUBMED 20083487  
REFERENCE 10617197  
AUTHORS 2 (bases 1 to 101365)  
Lin.X.  
TITLE Direct Submission  
COMMENT Submitted (09-MAR-2000) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
On Dec 17, 1999 this sequence version replaced gi:4581138.  
The sequence and annotation of chromosome 2 were merged from those  
of the individual clones on this chromosome after removing  
overlap. For detailed information, please see the TIGR web site  
(http://www.tigr.org/tdb/at/at.html).

Genes were identified by a combination of three methods: Gene  
prediction programs including GRATL  
(ftp://artur.epm.ornl.gov/pub/xgrat1), GeneFinder (Phil Green,  
University of Washington), Genscan (Chris Burge,  
http://genomic.stanford.edu/GENSCANW.html), and NetPlantGene  
(http://www.cbs.dtu.dk/services/NetGene2/), searches of the  
complete sequence against a peptide database and plant EST  
databases at TIGR, and manual curations based on those analyses.  
Annotated genes are named to indicate the level of evidence for  
their annotation. Genes with similarity to other proteins are named  
after the database hits. Genes without significant peptide  
similarity but with EST similarity are named as 'unknown' proteins.  
Genes without protein or EST similarity, that are predicted by two  
or more gene prediction programs over most of their length are  
annotated as 'hypothetical' proteins. Genes encoding tRNAs are  
predicted by tRNAscan-SE (Sean Eddy,  
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats were  
identified by RepeatMasker (Arian Smit,  
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Genes are  
numbered from the top to bottom of the chromosome.

We thank the GSHL/Mashu/ABI consortium for sequencing BAC clones  
F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone  
F1304, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards  
and Satoshi Tabata for helpful assistance. In addition, we would  
like to thank the TIGR Bioinformatics Department, especially Lixin  
Zhou, Hanif Khalek, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy  
Peterson, Michael Holmes, and Delwood Richardson for software and  
database support.

This work was supported by the National Science Foundation,  
Department of Energy and the US Department of Agriculture.

Address all correspondence to: attligr.org.

FEATURES  
Source location/Qualifiers  
1..101365  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="II"

misc-feature  
/note="Sequence from clone F2H17"  
CDS  
/note="F2H17.2, predicted by genscan"  
10101304..486,563..613,1357..1445,1521..1626,1770..1806,  
1886..2037,2371..2794,3184..3312,3391..3536,4096..4314,  
4639..4719,4850..4894,4936..5055)  
/gene="At2g36370"  
<304..>5055  
/gene="At2g36370"  
/note="F2H17.2, predicted by genscan"  
10101304..486,563..613,1357..1445,1521..1626,1770..1806,  
1886..2037,2371..2794,3184..3312,3391..3536,4096..4314,  
4639..4719,4850..4894,4936..5055)  
/gene="At2g36370"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="AAF18663.1"  
/db\_xref="GI:6598633"  
/translation="MALSLVEFVILVVKPCYQDDASSTIOEISIASIASWMS  
EILSYGVKRAHRTROYASGILVFPMPKSKRLQMLVSKDFYGVNVPYELMMCKYKH  
HLTVSEMDLVGLIIMLDAGRLSDLEPSSODNTINIMEOYRFLPLMFJAGRSKS  
HGFSPADOSTELVYKIKMNPSTCLVDSTODPDPDYVRLEFEIITDSCPOLNE  
ASLISILPNSTFANLNRKRSLESLKRLAAVLNITVSTLLELNFRELVEVLTVDSP  
ILPVQASVYSGVCYTSFSENSIALCATISMNEHMDINSVSNLQTLHMSKEGI  
SETSLNLIITHSOKMSICLRDTKVSQSLCEPSTLEALDISNTTESIDSLRSTVL  
TSIKHLOELATSYCGEISIOEFSKSMPLRLRLERYTRMMTNDLIVTOSCNLT  
ELISECGSTTENGVAISYCIATLEDFLRHNSGCIQKSFILDATIKRFFCLHLNNEQ  
FPLRLVSLDKWDKAGEGDPVEVSNT"  
3489..101164  
/note="Sequence from clone F1011"  
complement(9283..9377)  
/rpt\_family="Y(A)n"  
10101109..10507,10609..10729,10984..11157,11243..11533,  
11613..11703,11793..12693,12781..12941,13030..13133,  
13216..13381,13466..13543,13626..13735,13825..14157,  
14244..14534,14613..14696,14771..14904,14986..15213,  
15298..15469,15550..15804,15894..16203)  
/gene="At2g36380"  
<10119..16203  
/gene="At2g36380"  
/note="F1011.1: related to multi drug resistance proteins  
and P-glycoproteins"  
10101109..10507,10609..10729,10984..11157,11243..11533,  
11613..11703,11793..12693,12781..12941,13030..13133,  
13216..13381,13466..13543,13626..13735,13825..14157,  
14244..14534,14613..14696,14771..14904,14986..15213,  
15298..15469,15550..15804,15894..16163)  
/gene="At2g36380"  
/codon\_start=1  
/product="putative ABC transporter"  
/protein\_id="AAD24623.1"  
/db\_xref="GI:4581139"  
/translation="MIGREDLVPTMSGSGLSGSTSHRSLAGAASKSPFDVYAPETD  
VFGSDRREDDVELRMALERLPTPYDRLRKMLPQTMWNGKIGLEDVYVLMARKEK  
KHLMELKFEEDNEKFLRLRERTDRVGIQVPIEVENYELSVGGVRSARLDP  
LFWNLNTEISTIGLHILPSKKRIETKIDISGIIKPSRMILLGSPSGKTTLLA  
LAGKLDITLQMSGRITTCYCHEREYEPQATCYISQHDHREEMVRESLDSGCLG  
VQTRQLTELSRREERAGIKRDPEDLAPKSIATISQDSTSLVTVYLVKLKLDICAD  
TLVGVNRRGISGGOKRRLTJTEEMVGAFTALFMBEISYGLDSMTTFOCKMRDLVH  
IADVMVSILOPAPETFLPDITLISGQIVYVSGSRNVEFEFMEGFKPERKGI  
ADPLEVSTSKKDOEYWNREHYVSVSHVSDGNSPHAOCLASPERVYDAKT  
HPAALVTKQYVGSNKDLPKACDREMLMKRSPYVYVPTVOITIMSILAMVYRTE  
MHVGVVDDQKFTGALFESLINMEGRKALFVYWRLLVFFRKQDFRTPPWARLP  
GFLKIPSLIESVIALTYTTLTFAPAPAAEFQQLAVFCVNOALSTFFALFALG  
RFEVANSNGTLLALVVEVLKSGFEFFEPYFMICIGALLGFTLVENCYITLAMY  
RMGSPNDRINAKTVYEVILKSRGFEFFEPYFMICIGALLGFTLVENCYITLAMY  
NPLGNSKATTVVEGDKHSGSGVSTSTSSGPKGKAVLPPOPILSPANNVY  
YVDMPAEMKAGVVEGDRLLDLRDNGARPRVYITLNYGSGAKTMDVYLRKTCG  
YVGSINISGIFRNQKTRFARVSGYCBQNDIHSBPYHVESLITYSMLRSLADIDTKR  
EMFVEVEMLVELKPLRNSIVGLPGVDISTQDKRLTIAVELVANPSTIIEFDEPFG  
LDARAALIVMTVRNVTVDGRVVCVTIHPQSDIRLIDIESPEDELLMRGGVITAGTGLG  
HSOKIYVEFEALIEGVKIKIDGYNPATMMDVYTPSMESQNSPDFOIPIANSSVNRNO  
ELIKELSTPPPSNDLYFRYTKAOPFSYGTQKACPKMKWMSNRYQYVNAIRLMMVVI

[illegible][illegible]

exon 4316..4503  
/gene="F17C15\_10"  
/number=2  
4504..4804  
/gene="F17C15\_10"  
/number=2  
4805..4845  
/gene="F17C15\_10"  
/number=3  
4846..4953  
/gene="F17C15\_10"  
/number=3  
4954..5111  
/gene="F17C15\_10"  
/number=4  
5112..5202  
/gene="F17C15\_10"  
/number=4  
5203..5408  
/gene="F17C15\_10"  
/number=5  
8066..9485  
/gene="F17C15\_20"  
join(8066..8249,8411..8685,8882..9188,9283..9485)  
/gene="F17C15\_20"  
/note="similarity to coil protein po22,  
microspore/pollen-specific, alfalfa, PIR:T09416"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB82925.1"  
/db\_xref="GI:7340645"  
/translation="MGSNQLRPRKVPKLFVFGSDYADTNGTKRDTPEMAPIPYGITFP  
GKSGRYCDGLATDLEKRVKAGSPYLKRTGRKGLKRGMPAFGSKMLDDSPNS  
PPRNTTAQVNEFLVDYLAGRYGDTTPSPVSLISTAGDYITIIDNRPAGLKLVE  
KVDNLRVNMIVLGGLEKRLAVTSLQPLSPYTSASSFSKSCNCSALVELHNKL  
LKKVALLNEQSRVKKKEQHFEIIDIHNAFTVMKNGSKRFRKNPMKSCBEGYCGRSS  
DGKXYLTCDDEPSFFWDVAHPQTQGWKRSIVSLGNPLDTLTKP"  
8066..8249  
/gene="F17C15\_20"  
/number=1  
8250..8410  
/gene="F17C15\_20"  
/number=1  
8411..8685  
/gene="F17C15\_20"  
/number=2  
8686..8881  
/gene="F17C15\_20"  
/number=2  
8882..9188  
/gene="F17C15\_20"  
/number=3  
9189..9282  
/gene="F17C15\_20"  
/number=3  
9283..9485  
/gene="F17C15\_20"  
/number=4  
10909..10972  
/gene="F17C15\_30"  
/number=1  
10909..13585  
/gene="F17C15\_30"  
join(10909..10972,11060..11269,12610..12887,12976..13282,  
13365..13585)  
/gene="F17C15\_30"  
/note="similarity to proline-rich protein Apg, Arabidopsis  
thaliana, PIR:S21961"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB82926.1"  
/db\_xref="GI:7340646"  
/translation="MDSLILKFLCFLIFLCTSLFGEINGEGSNQNHLLYPRPTKL

intron FVFGSDYADTGNIKKAFSSMKPEYITFPCKPAGRFSGDHVATDFLAKFVGIKSPIP  
YEMKDYAGKKRLQYGMNFAFGTGVFNFTQTEPLPNNITQIDIFQNIILTTGDIYYPELT  
SSVALYVAGNDYSNEFTALNRPASFPATFIRKOVNDQEPVNIIRIHALGCKRIAPDSIQ  
PLGCLPPTPTVTSRQCNETONALYNLNNLLOQVAAKLNKETKOSTFTIILDLNAPL  
TVFKKNGSNPGSTREFSPKPCVGVSEYVNCGSVDKGVKKYITCDNPKPAFFWMDGL  
HPTLEGWRSYVSVLRESLTLASLKA"  
10973..11059  
/gene="F17C15\_30"  
/number=1  
11060..11269  
/gene="F17C15\_30"  
/number=2  
11270..12609  
/gene="F17C15\_30"  
/number=2  
12610..12887  
/gene="F17C15\_30"  
/number=3  
12888..12975  
/gene="F17C15\_30"  
/number=3  
12976..13282  
/gene="F17C15\_30"  
/number=4  
13283..13364  
/gene="F17C15\_30"  
/number=4  
13365..13585  
/gene="F17C15\_30"  
/number=5  
1397..14095  
/gene="F17C15\_40"  
/number=1  
1397..17133  
/gene="F17C15\_40"  
join(13397..14095,14188..14279,14346..14443,14535..14681,  
14766..14871,14965..15169,15255..15563,15651..15751,  
15832..16077,16155..16477,16559..17133)  
/gene="F17C15\_40"  
/note="strong similarity to cucumislin precursor,  
muskmelon, PIR:A55800  
contains Prokaryotic membrane lipoprotein lipid attachment  
site AAS36-546:Serine proteases, subtilase family, active  
sites AAS41-551"  
/codon\_start=1  
/product="cucumislin precursor-like protein"  
/protein\_id="CAB82927.1"  
/db\_xref="GI:7340647"  
/translation="MVSNOVRVLFMLCFQCIYNNNAVIAATPEDENVPRKPYIYMGPAE  
NSLVAENHHRHLMTVIGDESKARELKITISGKINQFVARLPFHEAKLSREGYV  
SVFKTQROQLHTTRSMDFLGLVESKFKRSVGIESNIIVGVDITGIDVSPSFNDGVG  
PEPPAKMGKGVGNFTRCNNKVIQAKFEHIOSEGLPGEDDTADHDHGHSTHTSTII  
AGVSYSASLFGIANGTARGCVPSARIKAYKQWDSCTDMDLAPDEALSDGDVSI  
STSIGASLPEFEDPIAIGAFHAKMRGILITCSAGNNGPGLTEVTNLSAGGVGPGSTCEPRT  
IDRKRETVVKKLNGILASGISINQGNPRKKMPLTSSGASLNSAGGVGPGSTCEPRT  
LGEDVYMKRVYICENAGREGGNGGQGHVYRSLKCAVITYQLLEPTPMATSTLAGS  
YVFEEDGTRITEYINSTKNPOAVIKTITTKMLAPSISFSARKGFORISPHLKPDIS  
APGLNIIAASKLASVATGVYPPDNRRTPLFSINSGTSMACPHAAAAAAYKSFHPDWSPA  
AIKSALMTTATPMRIKGNELSYSGQINPRRAIHPQLVYDITEDAVALRFLCKEGYN  
STSIGLITGDNSTNTTKKEYENCENIKRGLSGDLNYPSLHKOVNSTEAKVSEVFRTV  
TNVGYGSTYVARVWAPKGLRVEVVPKMWSEPRERKRNPKVYIDGVMDENMKGIYSA  
SVEMDSSGHLVRSPIILFRSDNDYR"  
14096..14187  
/gene="F17C15\_40"  
/number=1  
14188..14279  
/gene="F17C15\_40"  
/number=2  
14280..14345  
/gene="F17C15\_40"  
/number=2  
14346..14443



PHGSRKVRKLRMDTPGKVDKSIAMVIFKSOVQAGELIPFGNIYUADPEEEKYVYVHOHPQCP  
PESKRIYESHGMSPEPKRINSYANFDEVPYPRIKRIGYUANOIOMEHSYVAFSGK  
HYMFEPASSBEPCEPDILKSLIDRAHLELGLIYUADLHSHSNLNFGLDINFCDOTDRH  
YFHGPGRGHHMMDSRLFNWGSWEYLRFLSLNAKWMLBEIKFDDGREDGVTSMTHTTH  
GLQMTFGTNGEYFGEFATDVAVYLMVLNLIHGLYPAVASIEDYDSGMPTEPIPAV  
DGGGFDPIRLHMAVADKMTIELLKSDSESRKGGIDVHTLTJNRWLEKCTVYAESHDQAL  
VGDGTIAFWLMDKMDYPMALDRPSTPEIRGIALHMKIRLVTYMGLEGGYELNMGNGN  
EGHEWMDIPFGPOTLPRGKVLPCNNNSYDGCRRRFDGADDPYLRNMOGCGDOMAO  
LEEWGEPTSEHOHVSRKHEGDKVIFERGGYLVAFNPNHNSKSKDYVROFSGSKGKXK  
VALDSDALFGGFSRLDHDYDFTTEHRPHNDRPSFSVYTSRIAYVALTE"

BASE COUNT 731 a 494 c 632 g 657 t

ORIGIN

|    |  |                 |                   |           |              |
|----|--|-----------------|-------------------|-----------|--------------|
|    | Query Match  | 12.6%           | Score 125.8       | DB 8:     | Length 2554; |
|    | Best Local Similarity  | 84.0%:          | Pred. No.2.7e-18: |           |              |
|    | Matches 142:   | Conservative 0: | Mismatches 27:    | Indels 0: | Gaps 0:      |
| OY | 598 gctgttcacagatcacatgcttactcaatttttttgaccacaagtacgcgttttgaacctcca  | 657             |                   |           |              |
| Db | 820 GCGAGCTTTGGGTGCCATGCTTACTAATTTCCTTTTTTGACCAAGTAGGCCTTTTGAACTCCA  | 879             |                   |           |              |
| OY | 658 gagagaactaaatccctcgtatcgatatagagacaatsgcttgtgttgcttgcttcataatgat | 717             |                   |           |              |
| Db | 880 GAGGACTTAATAATCCTTTATCATGATGAGACACATGAGCTGTGGTTCCTGTTCTTATGAT    | 939             |                   |           |              |
| OY | 718 atgtgtcatagtagtaattagatgccaatatttatatttagctgttttacgcttt          | 766             |                   |           |              |
| Db | 940 ATTGTTCTATGICATTCGTCAAATATATACCCTTGACGGTTCGAATGGTT               | 988             |                   |           |              |

|            |   |
|------------|---|
| RESULT     | 8   |
| LOCUS      | AF338432  |
| DEFINITION | 2726 bp mRNA linear PLN 27-MAR-2001<br>Triticum aestivum starch branching enzyme I1a variant mRNA,  |
| ACCESSION  | AF338432  |
| VERSION    | AF338432.1 GI:13447951  |
| KEYWORDS   |   |
| SOURCE     | bread wheat.  |
| ORGANISM   | Triticum aestivum   |
| REFERENCE  | Euhayrola; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;<br>Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;<br>Pooidae; Triticeae; Triticum.<br>1 (bases 1 to 2726)<br>Rahman,S., Regina,A., Li,Z., Mukai,Y., Yamamoto,M.,<br>Kosar-Hashemi,B., Abrahams,S. and Morell,M.K.<br>Comparison of Starch-Branching Enzyme Genes Reveals Evolutionary<br>Relationships Among Isoforms. Characterization of a Gene for<br>Starch-Branching Enzyme I1a from the Wheat D Genome Donor Aegilops<br>tauschii |
| TITLE      | Plant Physiol. 125 (3), 1314-1324 (2001)  |
| JOURNAL    | 21140316  |
| MEDLINE    | 11244112  |
| PUBMED     | 2 (bases 1 to 2726)   |
| REFERENCE  | Rahman,S., Regina,A., Li,Z., Abrahams,S. and Morell,M.K.<br>Direct Submission<br>Submitted (18-JAN-2001) Plant Industry, CSIRO, Clunies Ross,<br>Canberra, ACT 2601, Australia  |
| AUTHORS    | Location/Qualifiers   |
| JOURNAL    | 1..2726   |
| FEATURES   | /organism="Triticum aestivum"<br>/db_xref="taxon:4565"  |
| SOURCE     | 124..2430   |
| CDS        |   |

```

/product="starch branching enzyme Iia variant"
/protein_id="AKK6822.1"
/db_xref="GI:13447952"
/translation="MAMFPAAGATLGNARPPAAAPPELQIPEDIEBQTAENMTGCT
AEKESSEPPQGIYETITDVTYKKEVLEVEKRVVFKPDGGQIETIDFLKDFR
HLDIRSEIRIRAIADQHEGLFASGIEIKLTSRAEITIKEMAPGASALVA
DFNMNPNADMTKDDIGVWEIIPNNADGSPALPHGSRVIRINDTQGVADSDISAW

```

BASE COUNT 726 a 564 c 701 g 735 t  
ORIGIN

| Query Match           | 12.4%  | Score 124.2       | DB 8     | Length 2726 |
|-----------------------|--|-------------------|----------|-------------|
| Best Local Similarity | 83.4%  | Pred. No. 6.1e-18 |          |             |
| Matches 141           | Conservative 0   | Mismatches 28     | Indels 0 | Gaps 0      |
| QY 598                | gcttctccaggtaccacatgcttactcaattcttttgaccacaagtacgcttttggaaactcca | 657               |          |             |
|                       |  |                   |          |             |
| Db 1039               | gcgacctttgggtaccatgcttactaatTTTTTTTgcaccacagtagcccttttggaaactcca | 1098              |          |             |
|                       |  |                   |          |             |
| QY 658                | gaggaactaaaatccctgatcgtatagtagacacatgagcttggcttgcgttctctatggat   | 717               |          |             |
|                       |  |                   |          |             |
| Db 1099               | gaggaactTAAATATCCCTCATGATGATGAGACACATGACGTGGTTGCTGTTCTTATGAT     | 1158              |          |             |
|                       |  |                   |          |             |
| QY 718                | attgtctatagtgtaattatgccaatttaattttagcgttttctacgctt               | 766               |          |             |
|                       |  |                   |          |             |
| Db 1159               | attgtctatagtgtaattatgccaatttaattttagcgttttctacgctt               | 1207              |          |             |
|                       |  |                   |          |             |

```
RESULT      9  
AF286319   AF286319       2970 bp    mRNA     linear    PLN 30-OCT-2000  
DEFINITION Triticum aestivum starch branching enzyme 2 (Sbe2) mRNA, complete cds.  
ACCESSION  AF286319  
VERSION     AF286319.1   GI:11037533  
KEYWORDS  
SOURCE  
ORGANISM    bread wheat.  
            Triticum aestivum  
REFERENCE   Emaripota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
            Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;  
            Poideae; Triticeae; Triticum.  
AUTHORS     1 (bases 1 to 2970)  
            McCue,K.F., Hurkman,W.J., Tanaka,C.K. and Anderson,O.D.  
TITLE        Starch Branching Enzymes Sbe1 and Sbe2 from Wheat (Triticum  
            aestivum cv. Cheyenne): Molecular Characterization, Developmental  
            Expression, and Homolog Assignment by Differential PCR  
            Unpublished  
JOURNAL      2 (bases 1 to 2970)  
REFERENCE    McCue,K.F. and Anderson,O.D.  
AUTHORS      Direct Submission  
TITLE        Submitted (11-JUL-2000) United States Department of Agriculture,  
            Agricultural Research Service, 800 Buchanan Street, Albany, CA  
            94710-1105, USA  
FEATURES  
source       Location/Qualifiers  
             1..2970-  
             /organism="Triticum aestivum"  
             /cultivar="Cheyenne"  
             /db_xref="taxon:4565"  
             /tissue_type="endosperm"  
             1..2970  
gene         /gene="Sbe2"
```

```

/ecf.number="2.4.1.81"
/note="glucosyltransferase: alpha-1,4-glucan:
alpha-1,4-glucan-6-glucosyltransferase"
/codon_start=1
/product="starch branching enzyme 2"
/protein_id="AG27623.1"
/db_xref="GI:11037534"

```

|    | Query Match           | 12.4%   | Score 124.2     | DB 8     | Length 2970 |
|----|-----------------------|---|-----------------|----------|-------------|
|    | Best Local Similarity | 83.4%   | Pred. NO. 6e-18 |          |             |
|    | Matches 141           | Conservative 0  | Mismatches 28   | Indels 0 | Gaps 0      |
| QY | 598                   | gctgttcacagtlacacatgtaactaatttttttgaccacagtcgcttttgaactcca      | 657             |          |             |
|    |                       |   |                 |          |             |
| Db | 1229                  | gcgacctttgggtaccatgtnactaatTTTTTTTGACCAAGAGCCGTTTGGAACTCCA      | 1288            |          |             |
|    |                       |   |                 |          |             |
| QY | 658                   | gaggaacttaaaatccttgatcgatagagacacatgagcttggttctgttcttaattgatt   | 717             |          |             |
|    |                       |   |                 |          |             |
| Db | 1289                  | GAGGACCTTAAAAATCCCGTATCGATGCAATGAGACACATGAGCTGTGGTTCTGTCTTATGAT | 1348            |          |             |
|    |                       |   |                 |          |             |
| QY | 718                   | attgttcagtagtaattagccaatttaattttagctgttttaacgctt                | 766             |          |             |
|    |                       |   |                 |          |             |
| Db | 1349                  | ATTGTTCTATAGTCATTCATCAATATATATCCCTTGACGGCTGTGAATAGGTT           | 1397            |          |             |
|    |                       |   |                 |          |             |

|             |  |
|-------------|--|
| RESULT      | 10   |
| LOCUS       | TASBA2   |
| DEFINITION  | T.aestivum mRNA for starch branching enzyme II.  |
| ACCESSION   | X11282   |
| VERSION     | X11282.1 GI:1885343  |
| KEYWORDS    | 1,4-alpha-glucan branching enzyme; sbe2 gene; starch branching enzyme II.  |
| SOURCE      | bread wheat.   |
| ORGANISM    | Triticum aestivum  |
| REFERENCE   | Eumaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.            |
| AUTHORS     | 1 (bases 1 to 2970) Nair,R.B., Baga,M., Scoles,G.J., Kartha,K.K. and Chibbar,R.N.  |
| TITLE       | Isolation, characterization and expression analysis of a starch branching enzyme II cDNA from wheat  |
| JOURNAL     | Plant Sci. In press  |
| REFERENCE   | 2 (bases 1 to 2970)  |
| AUTHORS     | Chibbar,R.N.   |
| TITLE       | Direct Submission  |
| JOURNAL     | Submitted (14-FEB-1997) R.N. Chibbar, Plant Biotechnology Institute, National Research Council of Canada, 110 Gymnasium Place, Saskatoon, Saskatchewan S7N 0W9, CANADA |
| FEATURES    | Location/Qualifiers  |
| Source      | 1..2970  |
|             | /organism="Triticum aestivum"  |
|             | /cultivar="Fielder"  |
|             | /db_xref="taxon:4565"  |
|             | /clone="PRN33"   |
|             | /dev_stage="kernels 12 days post anthesis"   |
|             | 151..2948  |
| gene        | /gene="sbe2"   |
|             | 151..312   |
| sig_peptide | /gene="sbe2"   |
|             | 151..2622  |
| CDS         | /gene="sbe2"   |

[illegible]

| Query Match           | 12.4%   | Score 124.2  | DB 8          | Length 2970     |
|-----------------------|---|--|---------------|-----------------|
| Best Local Similarity | 83.4%   | Pred. No. 6e-18  |               |                 |
| Matches 141           | Conservative  | 0  | Mismatches 28 | Indels 0        |
|                       |   |  |               | Gaps 0          |
| Oy                    | 598   | gctttccaggtaccacgtacgtaccatcttttttgaccacgaatagccgtttggaactcca    | 657           |                 |
|                       |   |  |               |                 |
| Db                    | 1231  | gcgacctttgggtraccancttacttatttttttgacaccagaagccgtttggaaactcca    | 1290          |                 |
| Oy                    | 658   | gaggaacttaaaatcccttgatcgatagagacatagagcttggcttgcgtttcttaaggat    | 717           |                 |
|                       |   |  |               |                 |
| Db                    | 1291  | gagagacttataaaatccctgatcgatagagacacatagacgttggcttgcgtttcttattgat | 1350          |                 |
| Oy                    | 718   | attgttcataagtgtaattagtcocaatttaatttagctgttttactgtttt             | 766           |                 |
|                       |   |  |               |                 |
| Db                    | 1351  | attgttcataagtcattcatccaatattatcccttgacggcgttgaaatggtt            | 1399          |                 |
| RESULT                | 11  |  |               |                 |
|                       | AF072725  | 23449 bp   | DNA           | linear          |
| LOCUS                 | AF072725  |  |               | PLN 02-SEP-1998 |
| DEFINITION            | Zea mays starch branching enzyme IIb (ze) gene, complete cds.       |  |               |                 |
| ACCESSION             | AF072725  |  |               |                 |
| VERSION               | AF072725.1  |  |               | GI:3511235      |
| KEYWORDS              | .   |  |               |                 |
| SOURCE                | Zea mays.   |  |               |                 |
| ORGANISM              | Zea mays  |  |               |                 |
|                       | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  |  |               |                 |
|                       | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC     |  |               |                 |
|                       | clade; Panicoideae; Andropogoneae; Zea.                             |  |               |                 |
| REFERENCE             | 1 (bases 1 to 23449)  |  |               |                 |
| AUTHORS               | Kim, K.-N., Fisher, D.K., Gao, M. and Gultinan, M.J.                |  |               |                 |
| TITLE                 | Molecular cloning and characterization of the amylose-extender gene |  |               |                 |
|                       | encoding starch branching enzyme IIb in maize                       |  |               |                 |
|                       | unpublished   |  |               |                 |
| JOURNAL               | 2 (bases 1 to 23449)  |  |               |                 |
| REFERENCE             | Kim, K.-N., Fisher, D.K., Gao, M. and Gultinan, M.J.                |  |               |                 |
| TITLE                 | Direct Submission   |  |               |                 |
| JOURNAL               | Submitted (16-JUN-1998) The Biotechnology Institute, Pennsylvania   |  |               |                 |

FEATURES State University, 302 Marlik, University Park, PA 16802, USA

SOURCE 1. 23449 /organism="Zea mays" /cultivar="B73" /db\_xref="taxon:4577"

Promoter 1. 2964 /gene="ae"

gene 1. 19879 /gene="ae"

TATA\_signal 2933. 2937 /gene="ae"

MRNA join(2965. 3176,3283. 3428,3673. 3827,4914. 5012, 5089. 5131,5328. 5387,5887. 5967,6049. 6165,6733. 6816, 7592. 7713,8465. 8584,12605. 12734,12821. 12931, 13080. 13208,16260. 16363,17236. 17380,17838. 17985, 18130. 18230,18457. 18534,18801. 18956,19405. 19479, 19576. 19879) /gene="ae"

5'UTR /product="starch branching enzyme IIB"

CDS 2965. 3064 /gene="ae"

join(3065. 3176,3283. 3428,3673. 3827,4914. 5012, 5089. 5131,5328. 5387,5887. 5967,6049. 6165,6733. 6816, 7592. 7713,8465. 8584,12605. 12734,12821. 12931, 13080. 13208,16260. 16363,17236. 17380,17838. 17985, 18130. 18230,18457. 18534,18801. 18956,19405. 19479, 19576. 19879) /gene="ae"

/note="SBE1B"

/codon\_start=1

/product="starch branching enzyme IIB"

/protein\_id="AAC33764.1"

/db\_xref="GI:3511236"

/translation="MAFVSGAVUGGAVRDLRTGGEGSLVFRHTGLFLTGARVGC SGTGAMAAAAARAAVAVPGEENDGLASRADSAQFQSEDELEVPDISSETTCAGVAD AQLNRVAVVPPSPDGOKLEQIDPMLOQYKYLEYRSLYRIRSDIDHEGGLAFA SRSYEGFENRGAEGITTYREMAFGAFSALVGDENNMDNADRMKSKEGWEIFLPNN ADGSPIDPHSGRVAVRMDTPSGIKDSIPAMIKYSQVARGELPYGITYDDPEEYVVF RHAQPKRKSLRIETHYGASSPEPKITTYNFRDEVLPRIKKICTYNAVOIMAOEHS YGSGYHVNTEFASSRFETPEELKSLIDRAHEGLLVLDVYHSHASSNTLDDNG FGDGTHYFHSRGRGHMMWDSRLFNYNMWELEFRLLSNAMWLEEFKDFRFGVAT SMATYHGLQVTFGENFENYEGFATDVAVVYLMVNLIGLYEAVTYIGEDVSGMP TEALPVHGAGVGFDRMMAVADKWDILDKOSDEMKWGDIVHTLTNRNMLEKCVTA ESHQALVGDCTIAFMIDKMDYPMALDREPTIDGIALHKMRITITMGLGEGY LNRMGNEGRPEWIDFRPGPRLPSGKATPGNNSTYDCRRRFDLGADYLRHGMQF FDQAMQHLQKYEFTSDHOYISRKHEEDKVIIVEKGDLVEFNFHCHNSYFDRIGC RKPQYKVVLDSDAGLFGGFSRIHHAHEFTADCSDHNRPSYSFVYTPSRICVYAPV E"

3'UTR 19660. 19879 /gene="ae"

polyA\_signal 19843. 19849 /gene="ae"

polyA\_site 19879 /gene="ae"

BASE COUNT 6409 a 4892 c 4904 g 7244 t

ORIGIN

Query Match 11.1%; Score 111.6; DB 8; Length 23449; Best Local Similarity 81.6%; Pred. No. 3.4e-15; Matches 129; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 309 tgcacgagcagcgaagataatcattcattcattttaggaagtggtttgcgaagaat 368

Db 7585 TGTATTGGAACCGAAGATTAACCATATGTAACTTAGGAGTGAAGTCTCCCAAGAAT 7644

QY 369 taaagccttgatacaagcagtcagataatgcaatccagcagcattcattactatgc 428

Db 7645 AAAAAAAGCTTGATACAAATGCAATTAATGCAATCCAGACACTCATATTATATG 7704

QY 429 aagcttgggtattcacacatcattttttctgtat 466

Db 7705 AAGCTTGGGTAATTCAGATCCAGTTTGTTCGCT 7742

RESULT 12

LOCUS E14724 2364 bp DNA linear PAT 28-JUL-1999

DEFINITION Rice mRNA for branching enzyme-4 mature peptide.

ACCESSION E14724

VERSION E14724.1 GI:5709407

KEYWORDS JP 1998004970-A/2.

SOURCE Oryza sativa.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 2364) Baba,T., Kawasaki,T. and Ichikawa,N. NEW RICE GENE FOR STARCH-BRANCHING ENZYME Patent: JP 1998004970-A 2 13-JAN-1998; MITSUI GYOSAI SHOKUBUTSU BIO KENYUSHO:KK, MITSUI PETROCHEM IND LTD

COMMENT OS Oryza sativa (rice) PN JP 1998004970-A/2 PD 13-JAN-1998 PF 24-JUN-1996 JP 1996162983 PI BABA TADASHI, KAWASAKI TSUTOMU, ICHIKAWA NORIO PC C12N15/09,A01H5/00,C07H21/04,C12N5/10,C12N9/10,(C12N15/09, PC C12R1:91), (C12N5/10,C12R1:91), (C12N9/10,C12R1:19); CC strandedness: Double; CC topology: linear; FH key Location/Qualifiers

FT source 1. 2364 /organism="Oryza sativa" /cultivar="Nihonbare" /clone\_lib="Rise Immature Seed Lambda gIII FT

FT cDNA library' /clone="PRB41" /issue\_type="Immature Seed" FT mat\_peptide 1. 2364

FEATURES source 1. 2364 /organism="Oryza sativa" /db\_xref="taxon:4530"

BASE COUNT 670 a 454 c 600 g 640 t

ORIGIN

Query Match 10.9%; Score 109; DB 6; Length 2364; Best Local Similarity 83.2%; Pred. No. 1.6e-14; Matches 124; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 313 caggaaacggaagataatcattcattcattttaggaagtggtttgcgaagaattaa 372

Db 859 CCGAAGCCGAAGATTAACACATATGCTATTAGGAGTACGATACCAAGATTAA 918

QY 373 aggccttgatacaatgagtcagataatgccaatccaggaacatcattcattgcaagc 432

Db 919 AAGCTTGGGTACAAATGCTGTACAGATAAAGCAATCCAGAGCACTTATTACCAAGC 978

QY 433 ttgggtattcacacatcatttttc 461

Db 979 TTTGGTATCATGTTACTTCTTTC 1007

RESULT 13

LOCUS E14723 3015 bp DNA linear PAT 28-JUL-1999

DEFINITION Rice mRNA for branching enzyme-4, complete cds.

ACCESSION E14723.1 GI:5709406

VERSION E14723.1 GI:5709406

KEYWORDS JP 1998004970-A/1.

| SOURCE   | ORGANISM  | REFERENCE                            | AUTHORS                               | JOURNAL                                   | COMMENT         |
|--|---|--------------------------------------|---------------------------------------|---|-----------------|
| Oryza sativa.  | Oryza sativa  | 1 (bases 1 to 3015)                  | Baba,T., Kawasaki,T., and Ichikawa,N. | NEW RICE GENE FOR STARCH-BRANCHING ENZYME |                 |
| Patent: JP 1998004970-A                                      | 1   | 13-JAN-1998                          |                                       |   |                 |
| MITSUI GROSAL SHOKUBUTSU BIO KENKYUSHO,KK,                   | MITSUI PETROCHEM IND LTD  |                                      |                                       |   |                 |
| OS   | Oryza sativa (rice)   |                                      |                                       |   |                 |
| PN   | JP 1998004970-A/1   |                                      |                                       |   |                 |
| PD   | 13-JAN-1998   |                                      |                                       |   |                 |
| PF   | 24-JUN-1996   | JP 1996162963                        |                                       |   |                 |
| PI   | BABA TADASHI, KAWASAKI TSUTOMU, ICHIKAWA NORIO                    | PC                                   |                                       |   |                 |
| C12N15/09,A01H5/00,C07H21/04,C12N5/10,C12N9/10,C12N15/09, PC |   |                                      |                                       |   |                 |
| C12R1.911',  |   |                                      |                                       |   |                 |
| PC   | (C12N5/10,C12R1.91),(C12N9/10,C12R1.19);                          |                                      |                                       |   |                 |
| CC   | strandedness: Double;   |                                      |                                       |   |                 |
| CC   | topology: Linear;   |                                      |                                       |   |                 |
| PH   | Key   | Location/Qualifiers                  |                                       |   |                 |
| FT   | source  | 1..3015                              |                                       |   |                 |
| FT   |   | /organism='Oryza sativa'             |                                       |   |                 |
| FT   |   | /cultivar='Nihonbare'                |                                       |   |                 |
| FT   |   | /clone.lib='Rice Immature Seed       |                                       |   |                 |
| FT   |   | library'                             |                                       |   |                 |
| FT   |   | /tissue.type='Immature seed'         |                                       |   |                 |
| FT   |   | /clone='pRB41'                       |                                       |   |                 |
| FT   | CDS   | 129..2654                            |                                       |   |                 |
| FT   |   | /product='branching enzyme-4'        |                                       |   |                 |
| FT   |   | 129..267.                            |                                       |   |                 |
| FT   |   | Location/Qualifiers                  |                                       |   |                 |
| FT   |   | 1..3015                              |                                       |   |                 |
| FT   |   | /organism='Oryza sativa'             |                                       |   |                 |
| FT   |   | /db.xref='taxon:4530'                |                                       |   |                 |
| FT   |   | 796 a 606 c 819 g 794 t              |                                       |   |                 |
| BASE COUNT   |   |                                      |                                       |   |                 |
| ORIGIN   |   |                                      |                                       |   |                 |
| Query Match  |   | 10.9%; Score 109; DB 6; Length 3015; |                                       |   |                 |
| Best Local Similarity  |   | 83.2%; Pred. No. 1.6e-14;            |                                       |   |                 |
| Matches 124; Conservative                                    |   | 0; Mismatches 25; Indels 0; Gaps 0;  |                                       |   |                 |
| OY 313   | caggaaecccaagaataatcatatgataattttagaggatgaggtgtccagaagatataa      | 372                                  |                                       |   |                 |
| Db 1146  | CCGGAAACCGAAGATTAACACATATGCTTAATTTAGGATGAGGTCTCCACAAAGATTATA      | 1205                                 |                                       |   |                 |
| OY 373   | aggacttgataacaatgcagtcagataatgccaatccagagagcatcactactatcagaac     | 432                                  |                                       |   |                 |
| Db 1206  | AAGCTTGGGTCACATGCTGTACAGATTAAATGGCAATCCAGGAGACTCTTTATTACCAAGC     | 1265                                 |                                       |   |                 |
| OY 433   | tttggtatcacaacatcatttttttc 461                                    |                                      |                                       |   |                 |
| Db 1266  | TTTGGGTATCATGTACTAATCTTCTTGC 1294                                 |                                      |                                       |   |                 |
| RESULT 14  |   |                                      |                                       |   |                 |
| LOCUS  | AB023498  |                                      |                                       |   |                 |
| DEFINITION   | Oryza sativa mRNA for starch branching enzyme rbe4, complete cds. |                                      |                                       |   |                 |
| VERSION  | AB023498.1  | GI:5689137                           |                                       |   |                 |
| KEYWORDS   | starch branching enzyme rbe4.                                     |                                      |                                       |   |                 |
| SOURCE   | Oryza sativa  |                                      |                                       |   |                 |
| ORGANISM   | Oryza sativa  |                                      |                                       |   |                 |
| REFERENCE  | AB023498  | 3015 bp                              | mRNA                                  | linear                                    | PLN 04-AUG-1999 |
| AUTHORS  | Mizuno,K., Tachibana,M., Kobayashi,E., Kawasaki,T., Funane,K.,    |                                      |                                       |   |                 |
| TITLE  | Kobayashi,M. and Baba,T.  |                                      |                                       |   |                 |
|  | Molecular cloning and expression analysis of a novel member of    |                                      |                                       |   |                 |

| REFERENCE                 | JOURNAL   | starch branching enzyme isoform in developing rice seeds  |
|---------------------------|---|---|
| 2 (bases 1 to 3015)       | Unpublished (1999)  |   |
| Authors                   | Mizuno, K. and Baba, T.   |   |
| JOURNAL                   | Direct Submission   |   |
|                           | Submitted (09-FEB-1999)   | Kouchi Mizuno, University of Tsukuba, Institute of Agricultural and Forest Engineering, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan |
|                           | E-mail: koumome@akura.cc.tsukuba.ac.jp, Tel: 81-298-524656, Fax: 81-298-55-2203   |   |
| FEATURES                  | Location/Qualifiers   |   |
| source                    | 1..3015   |   |
|                           | /organism="Oryza sativa"  |   |
|                           | /db_xref="taxon:4530"   |   |
|                           | /tissue_lib="developing seeds"  |   |
|                           | 129..2654   |   |
| gene                      | /gene="RBE4"  |   |
|                           | 129..2654   |   |
|                           | /gene="RBE4"  |   |
|                           | /codon_start=1  |   |
|                           | /product="starch branching enzyme rbe4"   |   |
|                           | /protein_id="BA82828.1"   |   |
|                           | /db_xref="GI:5689138"   |   |
|                           | /translation="MASFAVSGARLVNRRGGGGGGCGPANSRGVDLPSVLRKRD<br>EISFVGVSCAGAPGKVLVPGGSGDDLLLSAEPEVTEIDGPEEESDIPDNKVPKEE<br>EIPAVASAIKVAEDKLESSEVQIDIESEVPEIDADPEVIEDPRVIRPPGDDQ<br>KIYQIDPLEGFRHLDYRKYSEYRMAAIDQHEGGDLSRGYKGLGFRSAGITY<br>REMAQSGSAIVGDENNMNDMTNRNEGVWEISLPNNADGSPAIPIGSRVKIM<br>DTPSGVDSIPAKTKFAVQAPGEPYVNGIYVDPPEEKYVFOHPQKRPMSIRYESH<br>TGMSPPEKINTYANFDEVLPRIKKIGYNAVQIMAIQHSYASFGYHYTNFPASS<br>RFGSPEDKSLIDAHLEGLLVLMQDIVSHASNTLQNLNGFSDGTDRTHFGGRGH<br>MMDSRLFNNGSWEVRLRYLLSNARMWLEEKFDGFRDVGTSIMYTHGGLQVAFGNY<br>GEYFGEATDVDAVYLYLVNDLIHGILPEEVAIGEDVSGMPTFCIPVQDGGVGYDRL<br>HMAVDPKMIELKQSDDEYKMGKGDIVHTLTNRMSGEKVYAESHDQALVDKTIAYML<br>MDMDYFEMALDRSTPRIDRGIALHKMIRLYTMVGLGEGYLVNMGELNPFEDIDPP<br>RGQSPILNGSVLPNGNNSPDKCRPRRPLGADYIRVGLGEGYLVNMGELNPFEDIDPP<br>EHQYISKRHEDEKYLITERDDIVYVNFPHMSNSYEDRVGVCIKRGIKKIYLDSDGLF<br>GGFSRLDHADEYFADMPFHNDNRCPSVYTPRIYAVYALTED" |   |
| BASE COUNT                | 796 a 606 c 819 g 794 t   |   |
| ORIGIN                    |   |   |
| Query Match               | 10.9% Score 109; DB 8; Length 3015;   |   |
| Best Local Similarity     | 83.2% Pred. No. 1.6e-14;  |   |
| Matches 124; Conservative | 0; Mismatches 25; Indels 0; Gaps 0;   |   |
| QY                        | 313 cagggaaccgaagataaattcatalgtactaatttagggatggagtggtgccagaagataaa 372  |   |
|                           |   |   |
| Db                        | 1146 CCGGAACCGAAGATAAACAACATATGCTAATTTTGGAGGTGAGGTCTACCAAGAAATTAA 1205  |   |
|                           |   |   |
| QY                        | 373 aggccttgatcataatgcagtcagatataatggcaatccaggagcattactactatgcaagc 432  |   |
|                           |   |   |
| Db                        | 1206 AAGCTTGGGTCAATAGCTGTGACAGATTAATGGCAATCCAGGAGCACTTTATTAGCAAGC 1265  |   |
|                           |   |   |
| QY                        | 433 tttaggtatcacacaatcatcttttc 461  |   |
|                           |   |   |
| Db                        | 1266 TTGGGTATCATGTTACTAATCTCTTTC 1294   |   |
|                           |   |   |
| RESULT 15                 |   |   |
| ZM059948                  | 2M059948  |   |
| LOCUS                     | ZM059948  | 2795 bp mRNA linear PLN 21-AUG-1997   |
| DEFINITION                | Zea mays starch branching enzyme I1a (sbe2a) mRNA, partial cds.   |   |
| ACCESSION                 | U65948  |   |
| VERSION                   | U65948.1  | GI:2340107  |
| KEYWORDS                  |   |   |
| SOURCE                    | Zea mays.   |   |
| ORGANISM                  | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.  |   |
| REFERENCE                 | 1 (bases 1 to 2795)   |   |
| AUTHORS                   | Gao, M., Fisher, D.K., Kim, K.N., Shannon, J.C. and Guiltinan, M.J.   |   |







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 08:21:02 : Search time 406.92 Seconds  
(without alignments)  
4223.512 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_4500\_5500  
Perfect score: 1001  
Sequence: 1 tactagcttaagattccca.....tgatttaaccagttgaca 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 1001  | 100.0       | 11473  | 20    | AAK34650    |
| 2          | 965   | 96.4        | 11475  | 22    | AAH78338    |
| 3          | 124.2 | 12.4        | 2726   | 22    | AAH78337    |
| 4          | 112.2 | 11.2        | 8381   | 22    | AAH78343    |
| 5          | 111.6 | 11.1        | 23449  | 21    | AAZ35393    |
| 6          | 109   | 10.9        | 3015   | 19    | AAV05639    |
| 7          | 107   | 10.7        | 2307   | 19    | AAZ99938    |
| 8          | 105.4 | 10.5        | 2968   | 22    | AAH78342    |
| 9          | 100.8 | 10.1        | 2087   | 18    | AAH69737    |

|   |    |       |      |       |    |          |                    |
|---|----|-------|------|-------|----|----------|--------------------|
| C | 10 | 100.8 | 10.1 | 2165  | 18 | AAH69736 | Corn starch branch |
|   | 11 | 100.8 | 10.1 | 2640  | 19 | AAV70961 | DNA encoding maize |
|   | 12 | 100.8 | 10.1 | 2665  | 18 | AAH69729 | Plasmid pBE240 ins |
|   | 13 | 100.8 | 10.1 | 2725  | 19 | AAV29757 | Zea mays starch br |
|   | 14 | 97.6  | 9.8  | 2919  | 19 | AAQ73750 | Rice starch branch |
|   | 15 | 95.4  | 9.5  | 1919  | 19 | AAV38722 | CNA encoding star  |
|   | 16 | 91    | 9.1  | 3090  | 19 | AAV38720 | Full length cassav |
|   | 17 | 88.6  | 8.9  | 2715  | 21 | AAQ45939 | Arabidopsis thalia |
|   | 18 | 88    | 8.8  | 2529  | 17 | AAH42637 | Class A starch bra |
|   | 19 | 88    | 8.8  | 2531  | 17 | AAH17267 | Class A starch bra |
|   | 20 | 88    | 8.8  | 2576  | 17 | AAH42636 | Class A starch bra |
|   | 21 | 88    | 8.8  | 2578  | 17 | AAH42631 | Class A starch bra |
|   | 22 | 88    | 8.8  | 3003  | 17 | AAH42634 | Class A starch bra |
|   | 23 | 88    | 8.8  | 3033  | 17 | AAH42630 | Class A starch bra |
|   | 24 | 88    | 8.8  | 3231  | 17 | AAH42632 | Class A starch bra |
|   | 25 | 87.6  | 8.8  | 3074  | 18 | AAH6587  | Potato starch bran |
|   | 26 | 84.8  | 8.5  | 2975  | 17 | AAH42635 | Class A starch bra |
|   | 27 | 83    | 8.3  | 2913  | 19 | AAV38719 | Full length cassav |
|   | 28 | 71.2  | 7.1  | 1809  | 18 | AAH69753 | Corn starch branch |
|   | 29 | 71.2  | 7.1  | 1865  | 18 | AAH69748 | Corn starch branch |
|   | 30 | 71.2  | 7.1  | 2487  | 18 | AAH69747 | Corn starch branch |
|   | 31 | 71.2  | 7.1  | 2565  | 18 | AAH69752 | Corn starch branch |
|   | 32 | 71.2  | 7.1  | 2713  | 19 | AAV70962 | DNA encoding maize |
|   | 33 | 71.2  | 7.1  | 2763  | 19 | AAV29758 | Zea mays starch br |
|   | 34 | 71.2  | 7.1  | 2771  | 13 | AAQ24257 | Branching enzyme D |
|   | 35 | 71.2  | 7.1  | 2772  | 18 | AAH69740 | Plasmid pBE65 inse |
|   | 36 | 71.2  | 7.1  | 8119  | 21 | AAZ35392 | Maize starch bran  |
|   | 37 | 71    | 7.1  | 4563  | 22 | AAH30910 | Wheat starch bran  |
|   | 38 | 68.8  | 6.9  | 604   | 21 | AAH16018 | Human prostate can |
|   | 39 | 68.8  | 6.9  | 2899  | 22 | AAH02926 | Human shear stress |
|   | 40 | 68.8  | 6.9  | 3075  | 24 | AAH94880 | Human DNA sequence |
|   | 41 | 66.6  | 6.7  | 2687  | 20 | AAK34646 | WBE I-D4 cDNA seq  |
|   | 42 | 64.6  | 6.5  | 2733  | 15 | AAQ54674 | Rice starch branch |
|   | 43 | 64.6  | 6.5  | 2733  | 15 | AAQ62135 | Rice starch branch |
|   | 44 | 64.6  | 6.5  | 12151 | 15 | AAQ62137 | Rice starch branch |
|   | 45 | 63.4  | 6.3  | 11475 | 19 | AAV54979 | Potato starch bran |

## ALIGNMENTS

|             |  |                                   |
|-------------|--|-----------------------------------|
| RESULT 1    | AAK34650   | AAK34650 standard; DNA; 11473 BP. |
| XX          | AAK34650;  |                                   |
| AC          | AAK34650;  |                                   |
| XX          |  |                                   |
| 05-JUL-1999 | (first entry)  |                                   |
| XX          |  |                                   |
| DE          | Starch branching enzyme II (SBE II) gene sequence.                     |                                   |
| XX          |  |                                   |
| KW          | Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBF; GBS; |                                   |
| KW          | starch branching enzyme; starch soluble synthase; debranching enzyme;  |                                   |
| KW          | endosperm; wheat; barley; granule-bound synthase; glutenin; starch;    |                                   |
| KW          | grain softness protein I; bacterial isomylase; glycogen synthase;      |                                   |
| KW          | WBE I-D4 gene; ss.   |                                   |
| XX          |  |                                   |
| OS          | Triticum tauschii.   |                                   |
| XX          |  |                                   |
| PN          | WO9914314-A1.  |                                   |
| XX          |  |                                   |
| PD          | 25-MAR-1999.   |                                   |
| XX          |  |                                   |
| PF          | 11-SEP-1998; 98WO-AU00743.   |                                   |
| XX          |  |                                   |
| XX          | 20-MAR-1998; 98AU-0002509.   |                                   |
| PR          | 12-SEP-1997; 97AU-0009108.   |                                   |
| XX          |  |                                   |
| PA          | (CSIR ) COMMONWEALTH SCI & IND RES ORG.                                |                                   |
| PA          | (GOOD-) GOODMAN FIELDER LTD.   |                                   |
| PA          | (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.                                 |                                   |
| PA          | (AUSU ) UNIV AUSTRALIAN NAT.   |                                   |
| XX          |  |                                   |

PI L1 2, Morell M, Rahman S;  
XX  
DR WPI; 1999-229525/19.  
XX  
PT New isolated cereal plant enzyme genes used for, e.g. expression of  
PS antisense sequences of granule bound synthase  
XX  
PS Claim 8; Page 75-81; 171pp; English.  
XX  
CC The invention relates to a novel enzyme of starch biosynthetic pathway  
CC in a cereal plant, where the enzyme is selected from starch branching  
CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I and debranching  
CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or  
CC SBE I of rice or maize. The methods and products can be used for  
CC targeting expression specifically to the endosperm of the seeds of cereal  
CC plants such as wheat or barley. They can be used for the expression of  
CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low  
CC mol. wt. glutenin, grain softness protein I, bacterial isomylase,  
CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They  
CC can be used for modifying the characteristics of starch produced by a  
CC plant. The present sequence represents the SBE II gene sequence.  
CC  
XX  
SQ Sequence 11473 BP; 3096 A; 2263 C; 2423 G; 3691 T; 0 other;

Query Match 100.0%; Score 1001; DB 20; Length 11473;  
Best Local Similarity 100.0%; Pred. No. 3.6e-241;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tactagcttaagattccctcattggatgataagaataattcgatggagcgtccca 60  
DB 4510 tactagcttaagattccctcattggatgataagaataattcgatggagcgtccca 4569  
QY 61 agccattccctcattaaatgaagagagacaaaggaggaggaggagggttccc 120  
DB 4570 agccattccctcattaaatgaagagagacaaaggaggaggaggagggttccc 4629  
QY 121 ttcatattctgcgagcgatccaacttcattgtctcaggtgtaagctgacgacagg 180  
DB 4630 ttcatattctgcgagcgatccaacttcattgtctcaggtgtaagctgacgacagg 4689  
QY 181 atctcccatatgaagagatagatattcttctgtaactctggaactggaactgagct 240  
DB 4690 atctcccatatgaagagatagatattcttctgtaactctggaactggaactgagct 4749  
QY 241 tgaggcatcgctaatatatactatcatcacataactagagagatgacatgaatttag 300  
DB 4750 tgaggcatcgctaatatatactatcatcacataactagagagatgacatgaatttag 4809  
QY 301 tctgattctgcaggaacgaaagataaatcatatgcttaattttaaggatgaggtgttg 360  
DB 4810 tctgattctgcaggaacgaaagataaatcatatgcttaattttaaggatgaggtgttg 4869  
QY 361 ccaagaatataaaggcttgatgataatgacatgacgaataatgacgaatgacatca 420  
DB 4870 ccaagaatataaaggcttgatgataatgacatgacgaataatgacgaatgacatca 4929  
QY 421 tactatgaagcttgggtatcacacaatcatctttctgtaaaccttcaacca 480  
DB 4930 tactatgaagcttgggtatcacacaatcatctttctgtaaaccttcaacca 4989  
QY 481 tttagagctattacatccctaattgcttcagacataaaatatattgataataatcccttat 540  
DB 4990 tttagagctattacatccctaattgcttcagacataaaatatattgataataatcccttat 5049  
QY 541 tagatatataatatacaactactagattcttgaaaaagatcatttatctgtgtgct 600  
DB 5050 tagatatataatatacaactactagattcttgaaaaagatcatttatctgtgtgct 5109  
QY 601 tcttccaggtacatgttactaatctttctgacaaagtaacggtttggaaaccacag 660  
DB 5110 tcttccaggtacatgttactaatctttctgacaaagtaacggtttggaaaccacag 5169

QY 661 gactaaatcccttgatcatagatagacacatgagcttggttgcctgtcttaagataatt 720  
DB 5170 gactaaatcccttgatcatagatagacacatgagcttggttgcctgtcttaagataatt 5229  
QY 721 gtctatggtaattagttcccaatttaatttagcgttttaactgtttatctgtattcttaa 780  
DB 5230 gtctatggtaattagttcccaatttaatttagcgttttaactgtttatctgtattcttaa 5289  
QY 781 agggaaatcgagcaattatcatatcatgtccaagaagtcgagcgaagtcgaattgt 840  
DB 5290 agggaaatcgagcaattatcatatcatgtccaagaagtcgagcgaagtcgaattgt 5349  
QY 841 caaatctagatgagtcataaagaaatattggcaaaaactagagtgagcaaaaataaatttt 900  
DB 5350 caaatctagatgagtcataaagaaatattggcaaaaactagagtgagcaaaaataaatttt 5409  
QY 901 cccatcccaatgagcaggccctatcgccgaatatctttccatctatataatgtgcta 960  
DB 5410 cccatcccaatgagcaggccctatcgccgaatatctttccatctatataatgtgcta 5469  
QY 961 cgtactcttttctcagatgataatgaaccagttgaca 1001  
DB 5470 cgtactcttttctcagatgataatgaaccagttgaca 5510

RESULT 2  
AAH78338  
ID AAH78338 standard; cDNA; 11475 BP.  
XX  
AC AAH78338;  
XX  
DT 26-NOV-2001 (first entry)  
XX  
DE Nucleotide sequence of a starch branching enzyme designated F2.  
XX  
KW Wheat: starch branching enzyme; BEIib; SBE; transgenic plant;  
KW starch biosynthetic pathway; amylopectin; F2; amylose; ss.  
XX  
OS Aegilops tauschii.  
XX  
PN WO200162934-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 21-FEB-2001; 2001WO-AU00175.  
XX  
PR 21-FEB-2000; 2000AU-0005742.  
XX  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (GOOD-) GOODMAN FIELDER LTD.  
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.  
PI Morell M, Rahman S;  
PI  
XX  
DR WPI; 2001-570635/64.  
XX  
PT Nucleic acids encoding wheat starch branching enzyme IIB, useful for  
PT altering the amylose and amylopectin content of cereal plants, e.g.  
PT wheat and barley -  
XX  
XX  
PS Example 1; Fig 2; 103pp; English.  
XX  
CC The present sequence encodes a wheat starch branching enzyme of  
CC Aegilops tauschii, designated F2. A. tauschii is likely to be the  
CC ancestral D genome donor of wheat. Probes isolated from the present  
CC sequence were used to identify type II starch branching enzymes (SBEs)  
CC in wheat, especially BEIib. The BEIib nucleic acids may be used to  
CC genetically transform cereal plants such as wheat or barley and for  
CC altering their nutritional content by modulating the starch  
CC biosynthetic pathway to vary levels of amylopectin and/or amylose  
CC produced in the plant.  
XX  
SQ Sequence 11475 BP; 3093 A; 2259 C; 2423 G; 3689 T; 11 other;

|                            |        |                     |           |               |
|----------------------------|--------|---------------------|-----------|---------------|
| Query Match                | 96.4%; | Score 965;          | DB 22;    | Length 11475; |
| Best Local Similarity      | 99.4%; | Pred. NO. 3.8e-232; |           |               |
| Matches 1000; Conservative | 0;     | Mismatches 1;       | Indels 5; | Gaps 3;       |

|    |      |  |      |
|----|------|--|------|
| QY | 1    | tactagcttgaagattcccaacttagagatgtaagaatatgtgattggagcgctccacga       | 60   |
| Db | 4502 | tactagcttgaagattcccaacttagagatgtaagaatatgtgattggagcgctccacga       | 4561 |
| QY | 61   | agccattcttaccttatttaatgagagagacgaagggggggggggggggggtctcc           | 120  |
| Db | 4562 | agccattcttaccttatttaatgagagagacgaagggggggggggggggggtctcc           | 4621 |
| QY | 121  | ttaattatcttgcgagcgatccaanaacttccattgcttcgaggtgtaoactgcagag         | 180  |
| Db | 4622 | ttaattatcttgcgagcgatccaanaacttccattgcttcgaggtgtaoactgcagag         | 4681 |
| QY | 181  | attccccaattgaagaagggtatagatgaattctctgtaaccacttggaaacttgagct        | 240  |
| Db | 4682 | attccccaattgaagaagggtatagatgaattctctgtaaccacttggaaacttgagct        | 4741 |
| QY | 241  | tgaagcaatcgctataatatactacatccaaactacttaaggatgcatctgaa-attta        | 299  |
| Db | 4742 | tgaagcaatcgctataatatactacatccaaactacttaaggatgcatctgaaattta         | 4801 |
| QY | 300  | gtgtgatctctgcacaggaacccgaagaatcaatcatatgcttaattttaaggatgagtg       | 359  |
| Db | 4802 | gtgtgatctctgcacaggaacccgaagaatcaatcatatgcttaattttaaggatgagtg       | 4861 |
| QY | 360  | gcccaagattcaaaagggcttggtatagaatgcagtcgacgaatgagcaatccaggaatc       | 419  |
| Db | 4862 | gcccaagattcaaaagggcttggtatagaatgcagtcgacgaatgagcaatccaggaatc       | 4921 |
| QY | 420  | atactatgcagaagcttgggtatcacacaacatccattttcttgtataac-tcttcacc        | 478  |
| Db | 4922 | atactatgcagaagcttgggtatcacacaacatccattttcttgtataacatttcacc         | 4981 |
| QY | 479  | catttggagctatttaactcctaattgcttaatgatgcacataaataatttggatataatccctt  | 538  |
| Db | 4982 | catttggagctatttaactcctaattgcttaatgatgcacataaataatttggatataatccctt  | 5041 |
| QY | 539  | attgaataatagtagacaacacacacacttagtatctg--aaagaatcatttatgtgt         | 595  |
| Db | 5042 | attgaataatagtagacaacacacacacttagtatctgaaatgaaanaagatcatttatgtgt    | 5101 |
| QY | 596  | tgcgtcttccaggtaccacatgattcaattttttgcaccaaagtagccggtttggaact        | 655  |
| Db | 5102 | tgcgtcttccaggtaccacatgattcaattttttgcaccaaagtagccggtttggaact        | 5161 |
| QY | 656  | cagaggagcttaaaatccttgaatcgatagaaggacatgattggttgcgtgtactaagg        | 715  |
| Db | 5162 | cagaggagcttaaaatccttgaatcgatagaaggacatgattggttgcgtgtactaagg        | 5221 |
| QY | 716  | ataatgtctcataggaatattgttccaaatttaatttgcgcgttttaactgtttacggtat      | 775  |
| Db | 5222 | ataatgtctcataggaatattgttccaaatttaatttgcgcgttttaactgtttacggtat      | 5281 |
| QY | 776  | tctaaagggaatattcaggccaattatgatatcatgttcaaaagctaagaagtgcgaaagtga    | 835  |
| Db | 5282 | tctaaagggaatattcaggccaattatgatatcatgttcaaaagctaagaagtgcgaaagtga    | 5341 |
| QY | 836  | aataccaatctagagtggtgcaataaggaanaattggcaaaactctagatgagtcgaanaataa   | 895  |
| Db | 5342 | aatgccaatctctagagtggtgcaataaggaanaattggcaaaactctagatgagtcgaanaataa | 5401 |
| QY | 896  | attttccatctcaaatatgycagggccctatgcgcggaataattttccatctataataatg      | 955  |
| Db | 5402 | attttccatctcaaatatgycagggccctatgcgcggaataattttccatctataataatg      | 5461 |
| QY | 956  | tgtcagtgactctctttttctcagatgtaattaaaccagttgaca 1001                 |      |
| Db | 5462 | tgtcagtgactctctttttctcagatgtaattaaaccagttgaca 5507                 |      |

|          |                                  |
|----------|----------------------------------|
| RESULT   | 3                                |
| AAH78337 |                                  |
| ID       | AAH78337 standard; cDNA; 2726 BP |

|             |            |  |       |             |    |        |        |      |
|-------------|------------|--|-------|-------------|----|--------|--------|------|
| Query Match | Similarity | 12.4%  | Score | 124.2       | DB | 22     | Length | 2726 |
| Best Local  | Similarity | 83.4%  | Pred. | No. 2,1e-21 |    |        |        |      |
| Matches     | 141        | Conservative   | 0     | Mismatches  | 28 | Indels | 0      | Gaps |
| QY          | 598        | gcttctccaaggtacacatgttactaattttttgcaccagaagtagccgttttggaaatcca | 657   |             |    |        |        |      |
| Db          | 1039       | ggagagcttgggtacacatttactaattttttgcaccagaagtagccgttttggaaatcca  | 1098  |             |    |        |        |      |
| QY          | 658        | gagagacttaaaatccctttgcatgcatgagacacatggactgtgtgttgccttcttatgat | 717   |             |    |        |        |      |
| Db          | 1099       | gagagacttaaaatccctttgcatgcatgagacacatggactgtgtgttgccttcttatgat | 1158  |             |    |        |        |      |
| QY          | 718        | attgttcataaggttaatttagtccaatttaattttagctgttttactgtttt          | 766   |             |    |        |        |      |
| Db          | 1159       | attgttcataaggttaatttagtccaatttaattttagctgtgttttactgtttt        | 1207  |             |    |        |        |      |
| RESULT      | 4          |  |       |             |    |        |        |      |
| AAH78343    |            |  |       |             |    |        |        |      |
| ID          | AAH78343   | standard   | DNA   | 8381        | BP |        |        |      |
| XX          |            |  |       |             |    |        |        |      |
| AC          | AAH78343   |  |       |             |    |        |        |      |
| XX          |            |  |       |             |    |        |        |      |



```

FT      5388..5886
FT      /tag- w
FT      /number- 6
FT      5887..5967
FT      /tag- x
FT      /number- 7
FT      5968..6048
FT      /tag- y
FT      /number- 7
FT      6049..6165
FT      /tag- z
FT      /number- 8
FT      6166..6750
FT      /tag- aa
FT      /number- 8
FT      6751..6816
FT      /tag- ab
FT      /number- 9
FT      6817..7591
FT      /tag- ac
FT      /number- 9
FT      7592..7713
FT      /tag- ad
FT      /number- 10
FT      7714..8464
FT      /tag- ae
FT      /number- 10
FT      8465..8584
FT      /tag- af
FT      /number- 11
FT      8585..12604
FT      /tag- ag
FT      /number- 11
FT      12605..12734
FT      /tag- ah
FT      /number- 12
FT      12735..12820
FT      /tag- ai
FT      /number- 12
FT      12821..12931
FT      /tag- aj
FT      /number- 13
FT      12932..13079
FT      /tag- ak
FT      /number- 13
FT      13080..13208
FT      /tag- al
FT      /number- 14
FT      13209..16259
FT      /tag- am
FT      /number- 14
FT      16260..16363
FT      /tag- an
FT      /number- 15
FT      16364..17235
FT      /tag- ao
FT      /number- 15
FT      17236..17380
FT      /tag- ap
FT      /number- 16
FT      17381..17837
FT      /tag- ag
FT      /number- 16
FT      17838..17985
FT      /tag- ar
FT      /number- 17
FT      17986..18129
FT      /tag- as
FT      /number- 17
FT      18130..18230
FT      /tag- at
FT      /number- 18
FT      18231..18458

```

```

FT      /tag- au
FT      /number- 18
FT      18457..18534
FT      /tag- av
FT      /number- 19
FT      18535..18800
FT      /tag- aw
FT      /number- 19
FT      18801..18956
FT      /tag- ax
FT      /number- 20
FT      18957..19404
FT      /tag- ay
FT      /number- 20
FT      19405..19479
FT      /tag- az
FT      /number- 21
FT      19480..19575
FT      /tag- ba
FT      /number- 21
FT      19576..19659
FT      /tag- bb
FT      /number- 22
FT      19843..19849
FT      /tag- bc
FT      23089..23101
FT      /tag- bd

FT      misc-feature
FT      23089..23101
FT      /tag- bd

FT      W09964562-AZ.
FT      16-DEC-1999.
FT      11-JUN-1999; 99WO-US13266.
FT      12-JUN-1998; 98US-0089049.
FT      12-JUN-1998; 98US-0089050.
FT      (YPE-) UNIV PENNSYLVANIA STATE.
FT      Gulltlinan MJ, Kim K;
FT      WPI; 2000-116538/10.
FT      DR      XX
FT      PI      XX
FT      XX      XX
FT      PT      New gene regulatory sequences from plants used to provide resistance to
FT      PT      microbial path pathogens -
FT      XX

Query Match      11.1%; Score 111.6; DB 21; Length 23449;
Best Local Similarity 81.6%; Pred. NO.5.8e-18;
Matches 129; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY      309 tgcacaggaacgagaataatcatatgctaatttaggaatgagtggtgccaagaat 368
QY      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      7585 tgaatggaacggaagttaacacatctgtaaaccttagtgatgagtcctccaagaat 7644

QY      369 taaaagcttgatcaatcagtcagatgaatgccaatccagagacatcactatgc 428
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      7645 aaaaaaacttgatcaatcagtcagatgaatgccaatccagagacatcactatg 7704

QY      429 aagcttgggtatccacacatccatttttctgcat 466
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      7705 aagcttgggtatccacacatccatttttctgcat 7742

RESULT      6
ID      AAV05639
XX      AAV05639 standard; cDNA to mRNA; 3015 BP.
AC      AAV05639;
XX      AAV05639;
DT      01-MAY-1998 (first entry)
XX      Rice type IV starch branching enzyme cDNA.
DE

```

```

XX Rice: type IV starch branching enzyme; amylopectin synthesis; ds.
XX Oryza sativa.
XX Key Location/Qualifiers
XX CDS 129..2654
XX FT /*tag= a
XX FT 129..287
XX FT sig_peptide /*tag= b
XX FT mat_peptide 288..2651
XX FT /*tag= C
XX FT /product= type_IV_starch_branching_enzyme
XX JP10004970-A.
XX PD 13-JAN-1998.
XX PF 24-JUN-1996; 96JP-0162983.
XX PR 24-JUN-1996; 96JP-0162983.
XX PA (MITS-) MITSUI GYOSAI SHOKUBUNSU BIO KENKYUSHO.
XX PA (MITR ) MITSUI TOATSU CHEN INC.
XX DR WPI; 1998-133625/13.
XX DR P-PSDB; AAW41763.
XX PT Rice starch branching enzyme gene - synthesises amylopectin to yield
XX PT high quality starch
XX PS Claim 4; Pages 5-8; 13pp; Japanese.
XX CC The present sequence encodes the rice type IV starch branching
XX CC enzyme, which has the ability to synthesise amylopectin. The
XX CC quality of starch is improved by the use of the protein.
XX CC
XX SQ Sequence 3015 BP; 796 A; 606 C; 819 G; 794 T; 0 other;

Query Match 10.9%; Score 109; DB 19; Length 3015;
Best Local Similarity 83.2%; Pred.No.1.4e-17;
Matches 124; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 313 caggaccgaagataatcatatgcttaatttagggatgagtggttgccaagaat1aa 372
DB 1146 ccggaaccgaagataaacccatagttaatttagggatgagtggttgccaagaat1aa 1205
OY 373 aggtctgatacaatgcagtcagataatgcaatccaggagcatcattatgcaagc 432
DB 1206 aagcttggtacatgctgtacagataatgcaatccaggagcatcattatgcaagc 1265
OY 433 ttgggtattcacacatcattttttc 461
DB 1266 ttgggtattcatgttactactctttgc 1294

RESULT 7
AAZ99938
ID AAZ99938 standard; DNA; 2307 BP.
XX AC AAZ99938;
XX DT 25-JUL-2000 (first entry)
XX DE DNA encoding part of starch branching enzyme II clone 5A1.
XX KW Wheat; starch branching enzyme II; SBEII; SBEII-1; starch;
XX KW gelatinisation onset; transgenic plant; foodstuff; bakery product; ss.
XX OS Triticum aestivum.
XX FH Key Location/Qualifiers

```

```

FT CDS.
FT FT 1..2307
FT FT /*tag= a
FT FT /product= "starch branching enzyme II"
FT FT /transl_except= (pos: 2036..2038, aa: Xaa)
FT FT /transl_except= (pos: 2051..2053, aa: Xaa)
FT FT /transl_except= (pos: 2090..2092, aa: Xaa)
FT FT /transl_except= (pos: 2096..2098, aa: Xaa)
FT FT /transl_except= (pos: 2114..2116, aa: Xaa)
FT FT /transl_except= (pos: 2147..2149, aa: Xaa)
FT FT /transl_except= (pos: 2168..2170, aa: Xaa)
FT FT /transl_except= (pos: 2174..2176, aa: Xaa)
FT FT /transl_except= (pos: 2189..2191, aa: Xaa)
FT FT /transl_except= (pos: 2267..2269, aa: Xaa)
FT FT /transl_except= (pos: 2270..2272, aa: Xaa)
FT FT /transl_except= (pos: 2306..2307, aa: Asn)
FT FT /note= "Xaa is an unknown amino acid; the sequence
FT FT contains 10 internal stop codons, but no
FT FT termination codon"
FT FT
FT FT WO200015810-A1.
FT FT
FT FT 23-MAR-2000.
FT FT
FT FT 09-SEP-1999; 99WO-GB03011.
FT FT
FT FT 10-SEP-1998; 98EP-0307337.
FT FT
FT FT (PLAN-) PLANT BREEDING INT CAMBRIDGE LTD.
FT FT
FT FT Goldsbrough A, Colliwer S;
FT FT
FT FT WPI; 2000-271446/23.
FT FT
FT FT P-PSDB; AAY84408.
FT FT
FT FT DNA encoding wheat starch branching enzyme II isoforms, useful in
FT FT altering the characteristics of a plant, especially elevated starch
FT FT gelatinization onset and/or peak temperature -
FT FT
FT FT Disclosure; Fig 10; 197pp; English.
FT FT
FT FT The present sequence encodes a wheat starch branching enzyme II (SBEII)
FT FT of a novel subclass of known SBEII genes, designated SBEII-1. The
FT FT SBEII-1 genes are thought to have similar functional properties to the
FT FT maize SBEIIb gene. Starch branching enzymes catalyse the formation of the
FT FT alpha-1,6 linkages, creating branch points in the growing starch
FT FT molecule, via hydrolysis of an alpha-1,4 linkage followed by reattachment
FT FT of the released alpha-1,4-glucan chain to the same or another glucosyl
FT FT chain. SBEII polypeptides can be used to alter the characteristics of a
FT FT plant, in particular to alter starch so that it has an elevated
FT FT gelatinisation onset and/or peak temperature. Starch obtained from
FT FT transgenic plants is useful in the preparation or processing a foodstuff,
FT FT particularly bakery products.
FT FT
FT FT Sequence 2307 BP; 674 A; 440 C; 544 G; 635 T; 14 other;

Query Match 10.7%; Score 107; DB 21; Length 2307;
Best Local Similarity 71.0%; Pred.No.4.2e-17;
Matches 157; Conservative 0; Mismatches 60; Indels 4; Gaps 1;

OY 220 cctacttggaacttgagcttgaggcatgctaataatactatcatcaatacttag 279
DB 228 ccgaagaggagagatgtatgtatcaccatcccaactcaagacaa---aatcatg 283
OY 280 aggatcatctgaaatttagtgatcttgcaaggagcaagcaagataatcatatgct 339
DB 284 cggatataagaaacacatgttgcatgtagcccggaaccaagataacacacatagca 343
OY 340 aatttaggagatgagtggttgccaagaat1aaagcttgatgataatgcatgcaagata 399
DB 344 aacttcaggagatgagtggttcccaagaat1aaagacttggtataacatgcatgcaata 403
OY 400 atggaatccaggagatcattacatactatgcaagcttggta 440

```



Db 404 atggcaatccagagcaccatcactatggaagcttggta 444

# RESULT 8

AAH78342  
ID AAH78342 standard; cDNA; 2968 BP.

XX AAH78342;

DT 26-NOV-2001 (first entry)

XX Nucleotide sequence of wheat starch branching enzyme (BE1B).

KW Wheat; starch branching enzyme; BE1B; SBE; transgenic plant;  
KM starch biosynthetic pathway; amylopectin; amylose; ss.

XX Triticum sp.

XX WO200162934-A1.

XX 30-AUG-2001.

XX 21-FEB-2001; 2001WO-AU00175.

XX 21-FEB-2000; 2000AU-0005742.

XX (CSTR ) COMMONWEALTH SCI & IND RES ORG.

XX (GOOD-) GOODMAN FIELDER LTD.

XX (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

XX Morell M, Rahman S;

XX WPI; 2001-570635/64.

XX Nucleic acids encoding wheat starch branching enzyme IIB, useful for  
PT altering the amylose and amylopectin content of cereal plants, e.g.  
PT wheat and barley -

XX Claim 6; Fig 9; 103pp; English.

XX The present sequence encodes a wheat starch branching enzyme,  
CC designated BE1B. BE1B is a type II starch branching enzyme (SBE).  
CC The BE1B nucleic acids may be used to genetically transform cereal  
CC plants such as wheat or barley and for altering their nutritional  
CC content by modulating the starch biosynthetic pathway to vary levels  
CC of amylopectin and/or amylose produced in the plant.

XX Sequence 2968 BP; 784 A; 626 C; 796 G; 756 T; 6 other;

Query Match 10.5%; Score 105.4; DB 22; Length 2968;

Best Local Similarity 70.6%; Pred. No. 1.1e-16;

Matches 156; Conservative 0; Mismatches 61; Indels 4; Gaps 1;

QY 220 cctacttggaacttgagcttgagcgcctcaatataatacacaacttga 279

Db 1101 ccgaagaaggaggtatgtatcaacatcctcaacaaagcacaacaa 1156

QY 280 aggatcatggaatttagtgatcttgacaggaaccggaagataatctgt 339

Db 1157 cggatataaacaacatgttgcatgagtagcccggaacaaagatacaacatattga 1216

QY 340 aattttagggatgagtggtgccaagaatbaaaggcttgatatacatgagtcagata 399

Db 1217 aacttcaggagatgagtgctccaagaatbaaaggcttgatatacatgagtcagata 1276

QY 400 atgccaatccagagcattcatatgcaagcttggta 440

Db 1277 atgccaatccagagcattcatatgcaagcttggta 1317

RESULT 9

AA69737  
ID AA69737 standard; DNA; 2087 BP.

XX AA69737;

DT 10-SEP-1997 (first entry)

XX Corn starch branching enzyme IIB gene fragment in pBE96.

KW Starch branching enzyme IIB; SBE1B; corn; maize; antisense;  
KM amylopectin; transgenic plant; pBE96; ss.

XX Zea mays.

XX WO9722703-A2.

XX 26-JUN-1997.

XX 12-DEC-1996; 96WO-US19678.

XX 20-DEC-1995; 95US-0009113.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Broglie KE, Hubbard NL, Klein TM;

XX WPI; 1997-341694/31.

XX Transgenic corn in which grain derived starch fine structure is  
PT controlled - specifically branch chain distribution of amylopectin,  
PT useful in preparation of thickened foodstuff  
XX Example 4; Page 56-57; 92pp; English.

XX A cDNA insert (AA69737) in plasmid pBE96 comprises a 2.09 kb  
CC fragment of corn starch branching enzyme IIB (SBE1B) cDNA (see  
CC also AA69729), starting at the initiating ATG codon of the coding  
CC region and terminating 312 bp 5' of the translation termination  
CC codon. It was obtd. by site-specific mutagenesis of the SBE1B  
CC gene in plasmid pBE240 to generate an NcoI site at the ATG start  
CC site, and EcoRI digestion of the mutagenised plasmid. The 2.09 kb  
CC fragment was inserted, in sense orientation, between the zein  
CC promoter and 3' regions in vector pMI103 to produce pBE96.  
CC Co-suppression and antisense techniques can be used to inhibit  
CC SBE1B in transgenic plants and thereby alter starch properties.

XX Sequence 2087 BP; 565 A; 396 C; 553 G; 573 T; 0 other;

Query Match 10.1%; Score 100.8; DB 18; Length 2087;

Best Local Similarity 57.4%; Pred. No. 1.5e-15;

Matches 201; Conservative 0; Mismatches 147; Indels 2; Gaps 1;

QY 127 atcttcgagcagatcaaaaacttcattctgtctgagtgtagtactgaaggatcc 186

Db 707 atcttcacatcagaggaataaaggatcattcaactcctgatacaagatctcagtcagccc 766

QY 187 cattataaagagatata--agtaattcttgaacttctggaacttggaaacttgagtcctgag 244

Db 767 caggagaataacatatatgaggtattatgatctctcctgagaaggtaaagtatgt 826

QY 245 gcatcgctaataatactatcatcacaaataactaaggagatgagcttcgaattttagtg 304

Db 827 tcagagatcgcaaacctaaacgcaaaataacattgcgatatatgaacacatgcggaa 886

QY 305 atcttcacagaggaacggaataaattcatatgcttaatttagtgagtggttgccaa 364

Db 887 tgagtagcccggaaccggaataaacaacatattgaactttagggatgagtcctccaa 946

QY 365 gaattaaaggcttgataaatgacagtaaatgcaagtaaatgcaatccaggaattact 424

Db 947 gaataaaaacttgataaatgacagtaaatgcaagtaaatgcaatccaggaattact 1006



Db 772 caggagaataccatgatgatattatgatccctcctgaagagtaagtgtg 831  
Qy 245 gcatcgtaatatatactatcacatattcttagagagcatctyaaatttagtgt 304  
Db 832 tcaggcatcgcaacctaaacgacccaatcatctgcgatalatgaacacatgtcgaa 891  
Qy 305 atcttcacaggaacccgaataatcatatcttaattttagagatgagtggtcgcaa 364  
Db 892 tgaatgaccgggaacggaaataaacaacatgttaaactttagagatgagtggtcgcaa 951  
Qy 365 gaattaaagagcttgatataatgcatgcatgaatgcatgcatcaggagcatcatt 424  
Db 952 gaataaaaaaacttgatataatgcatgcatgaatgcatgcatcaggagcatcatt 1011  
Qy 425 atgcaagcttgggtatcacacatccatttttctgtatacactctt 474  
Db 1012 atggaagcttggatatacattgataactatttttgcgcaagtagtctg 1061

## RESULT 12

AA69729  
ID AAT69729 standard; DNA; 2665 BP.

AC AAT69729;

DT 10-SEP-1997 (first entry)

DE Plasmid PBE240 insert encoding corn starch branching enzyme IIb.

XX Starch branching enzyme IIb; SBEIIb; corn; maize; antisense;

KM amylopectin; transgenic plant; PBE240; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT CDS 79..2478

FT /\*tag- a

XX WO9722703-A2.

XX 26-JUN-1997.

XX 12-DEC-1996; 96WO-US19678.

XX 20-DEC-1995; 95US-0009113.

XX (DUPO ) DU PONT DE NEMOURS & CO E. I.

XX Broglie KE, Hubbard NL, Klein TM;

XX WPI; 1997-341694/31.

XX P-PSDB; AAW19212.

XX Transgenic corn in which grain derived starch fine structure is

XX controlled - specifically branch chain distribution of amylopectin,

XX useful in preparation of thickened foodstuff

XX Example 1; Page 50-53; 92pp; English.

XX A CDNA insert (AAT69729) in plasmid clone PBE240 comprises a 2.7 kb

XX EcoRI-XhoI fragment isolated from a corn cDNA library. It includes

XX an open reading frame encoding starch branching enzyme IIb (SBEIIb)

XX (AAW19212). The insert was used as a starting point in the assembly

XX of DNA constructs (see also AAT69730, AAT69736-37) designed to

XX achieve suppression of SBEIIb expression in transgenic corn plants,

XX and thereby to produce novel starches that have properties

Matches 201; Conservative 0; Mismatches 147; Indels 2; Gaps 1;

Qy 127 atctcgagcgatatacaaaacttccatttcttgagtgtaagtactgaggaactcc 186  
Db 785 atactccatcaggaataaaggatccattccagccttgatcaagtaactgagggcc 844  
Qy 187 cattatgaagaggaat--ggttaattcttgaacctacttgaaactgagtcgtgag 244  
Db 845 caggagaataccatgatgatgatttatatgactctctcctgaagaggtaaagtatgt 904  
Qy 245 gcatcgtaatatatactatcacatattcttagagagatgcatctgaaatttagtgt 304  
Db 905 tcaggcatcgcaacctaaacgacccaatcatctgcgatalatgaacacatgtcgaa 964  
Qy 305 atcttcacaggaacccgaataatcatatcttaattttagagatgagtggtcgcaa 364  
Db 965 tgaatgaccgggaacggaaataaacaacatgttaaactttagagatgagtggtcgcaa 1024  
Qy 365 gaattaaagagcttgatataatgcatgcatgaatgcatgcatcaggagcatcatt 424  
Db 1025 gaataaaaaaacttgatataatgcatgcatgaatgcatgcatcaggagcatcatt 1084  
Qy 425 atgcaagcttgggtatcacacatccatttttctgtatacactctt 474  
Db 1085 atggaagcttggatatacattgataactatttttgcgcaagtagtctg 1134

## RESULT 13

AAV29757  
ID AAV29757 standard; DNA; 2725 BP.

AC AAV29757;

DT 11-SEP-1998 (first entry)

DE Zea mays starch branching enzyme II gene.

XX SER; starch-encapsulating region; fusion vector;

KM starch branching enzyme II; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT CDS 91..2490

FT /\*tag- a

FT sig\_peptide /product= starch branching enzyme II

FT mat\_peptide /\*tag- b

FT /\*tag- c

XX WO9814601-A1.

XX 09-APR-1998.

XX 30-SEP-1997; 97WO-US17555.

XX 30-SEP-1996; 96US-0026855.

XX (EXSE-) EXSEED GENETICS LLC.

XX Guan H, Keeling P;

XX WPI; 1998-240100/21.

XX P-PSDB; AAW56489.

XX Hybrid polypeptide comprising starch-encapsulating region and

XX protein - useful for, e.g. producing protein(s) resistant to

XX degradation by stomach acids

XX Example 2; Page 44; 156pp; English.

Query Match 10.1%; Score 100.8; DB 18; Length 2665;

Best Local Similarity 57.4%; Pred. No. 1.6e-15;

The sequence is that of the starch branching enzyme II gene.



```

PT plants giving modified starch
XX
PS Claim 9; Fig 10; 67pp; English.
XX
CC The present sequence encodes starch branching enzyme (SBE) II. It
CC was isolated from cassava tubers. The products can be used for
CC producing plants having altered starch quantities and qualities.
CC They can also be used for producing altered plants such as cassava,
CC banana, potato, pea, tomato, maize, wheat, barley, oat, sweet potato
CC and rice plants.
XX
SQ Sequence 1919 BP; 543 A; 324 C; 467 G; 585 T; 0 other;

Query Match          9.5%; Score 95.4; DB 19; Length 1919;
Best Local Similarity 78.6%; Pred. No. 3.2e-14;
Matches 114; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 584 ttattgtgtgtgctgttcacaggtacatgttactaatttttgacccaagtggc 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 44 tttttttttttttagtttgggtacacatgttcacaaacttttgcacctagcagcc 103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 644 gtttggaaactccagaggaactaaacacctgacgatagagcacatgagctgttgc 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 104 gatttggaaactccctgatgatgttgaagctttaataagataaagctcatgagttaggcctgc 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 704 ttgtcttaagatattgttcataag 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 164 ttgtctcatgatattgttcataag 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: September 12, 2002, 08:21:51  
 Job time: 12268 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 08:12:27 ; Search time 92.79 Seconds  
(without alignments)  
2649.845 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_4500\_5500  
1001

Sequence: 1 tactagcttaagattccca.....tgtatcaaccagttgaca 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/lna/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/lna/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/lna/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/lna/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID                | Description       |
|------------|-------|-------------|--------|-------------------|-------------------|
| 1          | 100.8 | 10.1        | 2725   | US-08-941-445A-14 | Sequence 14, Appl |
| 2          | 87.6  | 8.8         | 3074   | US-09-087-277-1   | Sequence 1, Appl  |
| 3          | 86    | 8.6         | 1393   | US-09-087-277-3   | Sequence 3, Appl  |
| 4          | 71.2  | 7.1         | 2763   | US-08-941-445A-16 | Sequence 16, Appl |
| 5          | 63.4  | 6.3         | 11478  | US-08-981-803-29  | Sequence 29, Appl |
| 6          | 63.4  | 6.3         | 11478  | US-08-983-440-29  | Sequence 29, Appl |
| 7          | 61.8  | 6.2         | 2909   | US-08-104-158-1   | Sequence 1, Appl  |
| 8          | 61.8  | 6.2         | 3128   | US-08-716-449-1   | Sequence 1, Appl  |
| 9          | 54    | 5.4         | 807    | US-08-165-315D-1  | Sequence 1, Appl  |
| 10         | 54    | 5.4         | 807    | US-08-702-609A-2  | Sequence 2, Appl  |
| 11         | 54    | 5.4         | 846    | US-08-165-315D-2  | Sequence 2, Appl  |
| 12         | 54    | 5.4         | 1327   | US-08-165-315D-3  | Sequence 3, Appl  |
| 13         | 39.2  | 3.9         | 4467   | US-08-565-907A-1  | Sequence 1, Appl  |
| 14         | 39.2  | 3.9         | 4467   | US-08-910-551B-1  | Sequence 1, Appl  |
| 15         | 39.2  | 3.9         | 4467   | US-08-909-425A-1  | Sequence 1, Appl  |
| 16         | 38.6  | 3.9         | 8700   | US-08-392-625-16  | Sequence 16, Appl |
| 17         | 38.6  | 3.9         | 8700   | US-08-466-961A-16 | Sequence 16, Appl |
| 18         | 38.6  | 3.9         | 8700   | US-08-645-193B-18 | Sequence 18, Appl |
| 19         | 37.4  | 3.7         | 7218   | US-08-232-463-14  | Sequence 14, Appl |
| 20         | 37    | 3.7         | 84495  | US-09-797-906-3   | Sequence 3, Appl  |
| 21         | 35.8  | 3.6         | 1131   | US-08-951-924A-1  | Sequence 1, Appl  |
| 22         | 35.8  | 3.6         | 1131   | US-09-172-339-3   | Sequence 3, Appl  |
| 23         | 35.8  | 3.6         | 1131   | US-09-420-211-10  | Sequence 10, Appl |
| 24         | 35.8  | 3.6         | 2426   | US-08-528-026C-3  | Sequence 3, Appl  |
| 25         | 34.8  | 3.5         | 870    | PCT-US95-13658-3  | Sequence 3, Appl  |
| 26         | 34.2  | 3.4         | 1921   | US-08-557-128-11  | Sequence 11, Appl |
| 27         | 34.2  | 3.4         | 8133   | US-08-480-604A-5  | Sequence 5, Appl  |

28 34.2 3.4 8133 2 US-08-405-496A-5  
29 34.2 3.4 8133 4 US-08-935-136-5  
30 34 3.4 834 4 US-08-998-416-534  
31 33.8 3.4 398 3 US-08-714-918-5  
32 33.8 3.4 398 4 US-09-265-315-5  
33 33.8 3.4 398 4 US-09-265-315-5  
34 33.8 3.4 398 4 US-09-266-417-5  
35 33.8 3.4 4093 4 US-07-861-458C-5  
36 33.6 3.4 1185 2 US-08-179-557-14  
37 33.6 3.4 1680 4 US-09-242-690A-14  
38 33.4 3.3 312 4 US-09-242-690A-8  
39 33.4 3.3 662 4 US-08-998-416-185  
40 33.4 3.3 663 4 US-08-998-416-191  
41 33.4 3.3 665 4 US-08-998-416-937  
42 33.4 3.3 701 4 US-08-998-416-701  
43 33.4 3.3 711 4 US-08-998-416-786  
44 33.4 3.3 724 4 US-08-998-416-683  
45 33.4 3.3 732 4 US-08-998-416-1036

## ALIGNMENTS

RESULT 1  
US-08-941-445A-14  
Sequence 14, Application US/08941445A  
Patent No. 6107060  
GENERAL INFORMATION:  
APPLICANT: Keeling, Peter  
APPLICANT: Guan, Hanping  
TITLE OF INVENTION: Starch Encapsulation  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/941,445A  
FILING DATE: 30-SEP-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,855  
FILING DATE: 30-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 89-97  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2725 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 91..264  
FEATURE:  
NAME/KEY: mat\_peptide

LOCATION: 265..2487  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 91..2490  
US-08-941-445A-14

Query Match 10.1%; Score 100.8; DB 3; Length 2725;  
Best Local Similarity 57.4%; Pred. No. 5.9e-19;  
Matches 201; Conservative 0; Mismatches 147; Indels 2; Gaps 1;

QY 127 attctgcagcgtatcaaaactccattgtctgaggttgcgtactgcaggatctcc 186  
DB 797 ATACTCCATCAGGATTAAGAGATTCAATTCCAGCCTGGATCACTACAGTCCAGGCC 856  
QY 187 catatgaagaggat--agttattcttctgtacccactgtgaactgagcttgag 244  
DB 857 CAGAGAAATTCATATGATGGATTATTTATGATCTCTCGAAGAGSTAAAGTATGTGT 916  
QY 245 gcactgcataatactatcatcacatacttagagatgcatctgaatttagtgtg 304  
DB 917 TCAGGCATGCGCAACCTAAACGACCAAAATGATTCGCGATATATGAACAACATGTCCGAA 976  
QY 305 atcttcacaggaacccgaagataatcatctaatcttagagaggtgtgtccaa 364  
DB 977 TGACTACCCGCAACCAAGATTAACATATGTAACTTTAGGAGATGAAGTCTCCCAA 1036  
QY 365 gaataaagcgttgatataatgcagtcagataatgcagcagcagagatctact 424  
DB 1037 GAATTAATAAACTTGATATACATCATCTGCAATATATGCAATCCAAAGACATCATATT 1096  
QY 425 atgcaagcttggtatcacacaatccattttctgtatacactctt 474  
DB 1097 ATGGAAGCTTTGATACCATGTACTATTTTTCGCGCAAGTATGCGT 1146

RESULT 2  
US-09-087-277-1  
; Sequence 1, Application US/09087277B  
; Patent No. 6169226  
; GENERAL INFORMATION:  
; APPLICANT: EK, Bo  
; APPLICANT: KHOSNODI, Jamshid  
; APPLICANT: LARSSON, Clas-Tomas  
; APPLICANT: LARSSON, Hakan  
; APPLICANT: RASK, Lars  
; TITLE OF INVENTION: STARCH BRANCHING ENZYME II OF POTATO  
; FILE REFERENCE: 003300-486  
; CURRENT APPLICATION NUMBER: US/09/087, 277B  
; EARLIER APPLICATION NUMBER: PCT/SE96/01558  
; EARLIER FILING DATE: 1998-05-29  
; EARLIER FILING DATE: 1996-11-28  
; EARLIER APPLICATION NUMBER: SE 9504272-7  
; EARLIER FILING DATE: 1995-11-29  
; EARLIER APPLICATION NUMBER: SE 9601506-0  
; EARLIER FILING DATE: 1996-04-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3074  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:beli gene  
; OTHER INFORMATION: (branching enzyme II) from Solanum tuberosum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (189)..(2825)  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (189)..(332)  
; FEATURE:

NAME/KEY: mat\_peptide  
LOCATION: (333)..(2825)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (92)..(2156)  
OTHER INFORMATION: Nucleotides 92, 285, 1406, 1430, 1897 and 2156 are  
OTHER INFORMATION: n wherein n = A, C, G or T.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (285)..(287)  
OTHER INFORMATION: Amino acid -16 is Xaa wherein Xaa = Ile, Leu, Val  
OTHER INFORMATION: or Phe.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1404)..(1406)  
OTHER INFORMATION: Amino acid 358 is Xaa wherein Xaa = Leu or Phe.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1428)..(1430)  
OTHER INFORMATION: Amino acid 366 is Xaa wherein Xaa = Thr.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1896)..(1898)  
OTHER INFORMATION: Amino acid 522 is Xaa wherein Xaa = Tyr, Ser, Cys  
OTHER INFORMATION: or Phe.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (2154)..(2156)  
OTHER INFORMATION: Amino acid 608 is Xaa wherein Xaa = Pro.  
US-09-087-277-1

Query Match 8.8%; Score 87.6; DB 4; Length 3074;  
Best Local Similarity 73.0%; Pred. No. 3.1e-15;  
Matches 111; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 598 gctgttcacaggtaccatgttactaatcttttgcacaaagtacgcttttggaactcca 657  
DB 1374 gctagtttgttatcatcgtcacaaatttttngcaccaagcgcgttttggaancc 1433  
QY 658 gaggaactaaatcccttgatcgatagagacatgagcttggttctgtcttatgat 717  
DB 1434 gacgaccttagcttcttgatgataaagctcatgagctgaattgttctcatgac 1493  
QY 718 atgttcataagtaattagccaatttaatt 749  
DB 1494 atgttcacagcatgcatcaataataacttt 1525

RESULT 3  
US-09-087-277-3  
; Sequence 3, Application US/09087277B  
; Patent No. 6169226  
; GENERAL INFORMATION:  
; APPLICANT: EK, Bo  
; APPLICANT: KHOSNODI, Jamshid  
; APPLICANT: LARSSON, Clas-Tomas  
; APPLICANT: LARSSON, Hakan  
; APPLICANT: RASK, Lars  
; TITLE OF INVENTION: STARCH BRANCHING ENZYME II OF POTATO  
; FILE REFERENCE: 003300-486  
; CURRENT APPLICATION NUMBER: US/09/087, 277B  
; EARLIER APPLICATION NUMBER: PCT/SE96/01558  
; EARLIER FILING DATE: 1998-05-29  
; EARLIER FILING DATE: 1996-11-28  
; EARLIER APPLICATION NUMBER: SE 9504272-7  
; EARLIER FILING DATE: 1995-11-29  
; EARLIER APPLICATION NUMBER: SE 9601506-0  
; EARLIER FILING DATE: 1996-04-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1393



```

1 APPLICATION NUMBER: US 60/026,855
2 FILING DATE: 30-SEP-1996
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Winner, Ellen P
5 REGISTRATION NUMBER: 28,547
6 REFERENCE/DOCKET NUMBER: 89-97
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: (303) 499-8080
9 TELEFAX: (303) 499-8089
10 INFORMATION FOR SEQ ID NO: 16:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 2763 base pairs
13 TYPE: nucleic acid
14 STRANDEDNESS: single
15 TOPOLOGY: not relevant
16 MOLECULE TYPE: mRNA
17 HYPOTHEICAL: NO
18 ORIGINAL SOURCE:
19 ORGANISM: Zea mays
20 FEATURE:
21 NAME/KEY: transit_peptide
22 LOCATION: 2..190
23 FEATURE:
24 NAME/KEY: mat_peptide
25 LOCATION: 191..2467
26 FEATURE:
27 NAME/KEY: CDS
28 LOCATION: 2..2470

```

|    | Query Match | Similarity  | 7.1%   | Score 71.2:       | DB 3   | Length 2763: |
|----|-------------|---|--------|-------------------|--------|--------------|
|    | Best Local  | Similarity  | 69.3%: | Pred No. 1.2e-10: |        |              |
|    | Matches     | 97; Conservative  | 0;     | Mismatches 45;    | Indels | Gaps 0;      |
| Oy | 589         | tgtgttctgagcctgttcacagtacatgttaactaaattttttgcaccaagaagccgtttt     | 648    |                   |        |              |
| Db | 893         | TGCTACTATAGCTTCTTTGGGGGTACCATGTCGAATAATTCTTTGGCGGTGACAGCAGATCA    | 952    |                   |        |              |
| Oy | 649         | ggacctcacagaggactaaatccttgatcgatatgaagcaacatgatgcttgttctgttt      | 708    |                   |        |              |
| Db | 953         | GGCACACCACAGAGAGACCTCAAATATCTTTGTATAGAAGCCACACAGTGTTGGGTTGCCAGATT | 1012   |                   |        |              |
| Oy | 709         | ccttatgatattgttccaatag  | 728    |                   |        |              |
| Db | 1013        | CTGATGGATGTTGTCCATAG  | 1032   |                   |        |              |

```

RESULT      5
US-08-981-803-29
; Sequence 29, Application US/08981803
; Patent No. 6147279
; GENERAL INFORMATION:
; APPLICANT: POULSEN, PETER
; TITLE OF INVENTION: INHIBITION OF GENE EXPRESSION
; FILE REFERENCE: 674509-2004
; CURRENT APPLICATION NUMBER: US/08/981, 803
; CURRENT FILING DATE: 1997-04-17
; EARLIER APPLICATION NUMBER: PCT/EP96/03052
; EARLIER FILING DATE: 1996-07-12
; EARLIER APPLICATION NUMBER: 9514435.8
; EARLIER FILING DATE: 1995-07-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 29
; LENGTH: 11478
; TYPE: DNA
; ORGANISM: Solanum tuberosum
US-08-981-803-29

Query Match      6.3%; Score 63.4; DB 3; Length 11478
Best Local Similarity 70.2%; Pred. No. 3.3e-08;

```





US-08-702-609A-2

US-08-165-315D-2

853 +accataaagaaaattttccaaaaaactatgattttccaaaaataaaattttcc 902

Db 233 TGACATAAACAATAATTACCCATCATAAATGACACAAATAATTTCACC 184

RESULT 12

US-08-165-315D-3/C  
; Sequence 3, Application US/08165315D  
; Patent No. 5525716

GENERAL INFORMATION:

APPLICANT: Odd-Arne Olsen  
APPLICANT: Roger Kalla  
TITLE OF INVENTION: Promoter  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: St. Onge, Steward, Johnston & Reens  
STREET: 986 Bedford Street  
CITY: Stamford  
STATE: Connecticut  
COUNTRY: U.S.A.  
ZIP: 06905

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" 1.44 Mb diskette  
COMPUTER: IBM PC  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Word Processor  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/165,315D  
FILING DATE: 10 December 1993  
CLASSIFICATION: 800

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9324707.0  
FILING DATE: 2 December 1993

ATTORNEY/AGENT INFORMATION:

NAME: Mary M. Krinsky  
REGISTRATION NUMBER: 32423  
REFERENCE/DOCKET NUMBER: 2105-P0001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-324-6155  
TELEFAX: 201-327-1096  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1327  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
DESCRIPTION: gene  
FRAGMENT TYPE: gene  
FEATURE:  
NAME/KEY: Ltp2 gene  
US-08-165-315D-3

Query Match 5.4%; Score 54; DB 1; Length 1327;  
Best Local Similarity 68.2%; Pred. No. 6e-06;

Matches 75; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 793 gcaattatgatacttgcataaagctgaagtcgcaagatgcgaatctcaatctcaag 852

Db 293 GGAATGGAATTAATTGTTCAACCTAGAGTGGCAAAACGTGAATGTCAAAACCTAGAG 234

Qy 853 tggcataaggaataatggcaaaactagagtcgcaaaactaaatcttc 902

Db 233 TGACATAAACAATAATTACCCATCATAAATGACACAAATAATTTCACC 184

RESULT 13

US-08-565-907A-1/C  
; Sequence 1, Application US/08565907A  
; Patent No. 5814499

GENERAL INFORMATION:

APPLICANT: Sylvain Moineau, Barbara  
APPLICANT: J. Holler, Peter A. Vandenberg,  
APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.

APPLICANT: Kondo

TITLE OF INVENTION: DNA Encoding Phage  
TITLE OF INVENTION: Abortive Infection Protein

TITLE OF INVENTION: From Lactococcus  
TITLE OF INVENTION: lactis, and Method of use thereof

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ian C. McLeod  
STREET: 2190 Commons Parkway  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette 5.25 inch,  
MEDIUM TYPE: 360 kb storage  
COMPUTER: Acer  
OPERATING SYSTEM: MS-DOS (version 4)  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/565,907A  
FILING DATE: December 1, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: Quest 4.1-152  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
TELEX: No. 3814499e  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 4467  
TYPE: Nucleotide  
STRANDEDNESS: Single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
DESCRIPTION: No  
HYPOTHETICAL: No  
ANTI-SENSE: No  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:  
ORGANISM: Lactococcus lactis  
STRAIN:  
INDIVIDUAL ISOLATE: W1  
DEVELOPMENTAL STAGE: N/A  
HAPLOTYPE: N/A  
TISSUE TYPE: N/A  
CELL TYPE: Bacterium  
CELL LINE: N/A  
ORGANELLE: N/A  
IMMEDIATE SOURCE:  
LIBRARY: genomic  
CLONE: SMO-20  
POSITION IN GENOME: N/A

FEATURE:

NAME/KEY: phage abortive infection  
LOCATION: N/A  
IDENTIFICATION METHOD: sequencing  
OTHER INFORMATION: DNA encoding phage  
OTHER INFORMATION: resistance  
PUBLICATION INFORMATION: N/A

US-08-565-907A-1

Query Match 3.9%; Score 39.2; DB 1; Length 4467;  
Best Local Similarity 44.9%; Pred. No. 0.14;

Matches 149; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

```

QY 457 ttctgtatcaccttccaccatttgagactatcacatcgaatgctcagcacata 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2359 TATTGCTATATTTTAGTATCTACTCTTTTGTTCATCATGATTTGTTAATCGGA 2300
QY 517 aaatattgagataatcccttattagatataagatacaactacactagtagtcgaa 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2299 AAATACATTTTATAGCTCCCTTATCCCTAAGTTCCTTCAACCAATATATATA 2240
QY 577 aagatcatttattgttgttgctgtccaggtagacatgttactaattttttgacca 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2239 AATTGCTATTTCTTTAATCCACTTGCGTTGAATTAAGTAAATTTTCAAAAA 2180
QY 637 agtagccgttttgaaactccagagagactaaatcccttgatcgatagacatgagctt 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2179 AAGAAAAAATATCCGATTTCTCGATTTATCAACAAACGGAAATGTCAACTTCGTT 2120
QY 697 ggttgctgtcttcttgatgatatgttcataagtaattagccaatttaatttagctgt 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2119 TATTATCATTAATTAATTAAGTATATTTCTCGACAGATTAGATTAAATTCATTTAAAAAT 2060
QY 757 ttactgttatctgttattcttaagagaaat 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2059 CTGCTTTTATCTCTCAAAAGTAAACGGAAT 2028

```

## RESULT 14

US-08-910-551B-1/c

; Sequence 1, Application US/08910551B

; Patent No. 5910571

; GENERAL INFORMATION:

; APPLICANT: Sylvain Moineau, Barbara

; APPLICANT: J. Holler, Peter A. Vandenberg, R.

; APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.

; APPLICANT: Kondo

; TITLE OF INVENTION: DNA Encoding Phage

; TITLE OF INVENTION: Abortive Infection Protein

; TITLE OF INVENTION: From Lactococcus

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ian C. McLeod

; STREET: 2190 Commons Parkway

; CITY: Okemos

; STATE: Michigan

; COUNTRY: USA

; ZIP: 48864

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette 5.25 inch,

; MEDIUM TYPE: 360 Kb storage

; OPERATING SYSTEM: MS-DOS (version 4)

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/910,551B

; FILING DATE: August 11, 1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/565,907

; FILING DATE: December 1, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Ian C. McLeod

; REGISTRATION NUMBER: 20,931

; REFERENCE/DOCKET NUMBER: Quest 4.1-158

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (517) 347-4100

; TELEFAX: (517) 347-4103

; TELEX: No. 5910571e

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4467

; TYPE: Nucleotide

; STRANDEDNESS: Single

; TOPOLOGY: Linear

```

; MOLECULE TYPE: Genomic DNA
; DESCRIPTION: No
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: N/A
; ORIGINAL SOURCE:
; ORGANISM: Lactococcus lactis
; STRAIN:
; INDIVIDUAL ISOLATE: W1
; DEVELOPMENTAL STAGE: N/A
; HAPLOTYPE: N/A
; TISSUE TYPE: N/A
; CELL TYPE: bacterium
; CELL LINE: N/A
; ORGANELL: N/A
; IMMEDIATE SOURCE:
; LIBRARY: genomic
; CLONE: SMO-20
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: phage abortive infection
; LOCATION: N/A
; IDENTIFICATION METHOD: sequencing
; OTHER INFORMATION: DNA encoding phage
; OTHER INFORMATION: resistance
; PUBLICATION INFORMATION: N/A
US-08-910-551B-1

```

```

Query Match 3.9%; Score 39.2; DB 2; Length 4467;
Best local Similarity 44.9%; Pred. No. 0.14;
Matches 149; Conservative 0; Mismatches 183; Indels 0;

```

```

QY 457 ttctgtatcaccttccaccatttgagactatcacatcgaatgctcagcacata 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2359 TATTGCTATATTTTAGTATCTACTCTTTTGTTCATCATGATTTGTTAATCGGA 2300
QY 517 aaatattgagataatcccttattagatataagatacaactacactagtagtcgaa 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2299 AAATACATTTTATAGCTCCCTTATCCCTAAGTTCCTTCAACCAATATATATA 2240
QY 577 aagatcatttattgttgttgctgtccaggtagacatgttactaattttttgacca 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2239 AATTGCTTATTTCTTTAATCCACTTGCGTTGAATTAAGTAAATTTTCAAAAA 2180
QY 637 agtagccgttttgaaactccagagagactaaatcccttgatcgatagacatgagctt 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2179 AAGAAAAAATATCCGATTTCTCGATTTATCAACAAACGGAAATGTCAACTTCGTT 2120
QY 697 ggttgctgtcttcttgatgatatgttcataagtaattagccaatttaatttagctgt 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2119 TATTATCATTAATTAATTAAGTATATTTCTCGACAGATTAGATTAAATTCATTTAAAAAT 2060
QY 757 ttactgttatctgttattcttaagagaaat 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2059 CTGCTTTTATCTCTCAAAAGTAAACGGAAT 2028

```

## RESULT 15

US-08-909-425A-1/c

; Sequence 1, Application US/08909425A

; Patent No. 5928688

; GENERAL INFORMATION:

; APPLICANT: Sylvain Moineau, Barbara

; APPLICANT: J. Holler, Peter A. Vandenberg, R.

; APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.

; APPLICANT: Kondo

; TITLE OF INVENTION: DNA Encoding Phage

; TITLE OF INVENTION: Abortive Infection Protein

; TITLE OF INVENTION: From Lactococcus

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

```

ADDRESS: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 5.25 inch,
MEDIUM TYPE: 360 Kb storage
COMPUTER: Acer
OPERATING SYSTEM: MS-DOS (version 4)
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/909,425A
FILING DATE: August 11, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/565,907
FILING DATE: December 1, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5928688e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4467
TYPE: Nucleotide
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE:
DESCRIPTION: Genomic DNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N/A
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis
STRAIN:
INDIVIDUAL ISOLATE: W1
DEVELOPMENTAL STAGE: N/A
HAPLOTYPE: N/A
TISSUE TYPE: N/A
CELL TYPE: bacterium
CELL LINE: N/A
ORGANELLE: N/A
IMMEDIATE SOURCE:
LIBRARY: genomic
CLONE: SMQ-20
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: phage abortive infection
LOCATION: N/A
IDENTIFICATION METHOD: sequencing
OTHER INFORMATION: DNA encoding phage
OTHER INFORMATION: resistance
PUBLICATION INFORMATION: N/A
US-08-909-425A-1

```

```

Query Match 3.9%; Score 39.2; DB 2; Length 4467;
Best Local Similarity 44.9%; Pred. No. 0.14;
Matches 149; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

```

```

QY 457 ttcttgataactcttaccattgagcctatacatcccaatgcttcacacata 516
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2359 TATTGCTATATTATTTAGTATCTACTTTTGTTCATGATTTGTTAATACTGGGA 2300
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 517 aaatattgataaaccttataagataatagatacaactaacacttagctcga 576
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2299 AATACATTTTATAGCTCCCTTATTCCTTAATGTTCTTCATTCACACATATCTATA 2240
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY 577 aagatcatttattgttgtgctgtccaggctaccagctactaatttttgcacca 636
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2239 AATTGCTATTTCCTTAATCCACTGTGTTGGAATTAATTAATTTCAAAAA 2180
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 637 agtagccgttttggaaatccagagactaaatcccttgatcgatgacacagcgtt 696
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2179 AAGAAAAAATATCCGATTTACTCGATTTATCAACAAACGGGAATGTCACATTCGTTT 2120
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 697 ggttgctgttcttatagatalgttcatalagylaatagtcacatttaattagctgt 756
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2119 TATTATCATTAATAATTAAGTTATTTTCGACAGATTTAATTAATTAATAAATT 2060
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 757 ttactgtttactcgtglatcttaagaggaat 788
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2059 CTTCCTTTTCATTCACAAAAGTAACGGAAT 2028
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: September 12, 2002, 08:12:56  
Job time: 14833 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 08:10:10 : Search time 3207.55 Seconds  
(without alignments)  
4212.078 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_4500\_5500

Perfect score: 1001

Sequence: 1 tactagcttaagattccca.....tgtattaccagcttgaca 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estcmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_pln:\*  
15: em\_gss\_inv:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 125.8 | 12.6        | 678    | 9     | AV939010 AV939010  |
| 2          | 104.8 | 10.5        | 375    | 10    | BG240463 OVL_30.D0 |
| 3          | 100.8 | 10.1        | 639    | 9     | AM065909 687002G09 |
| 4          | 92.2  | 9.2         | 386    | 12    | CNS00SBA           |
| 5          | 89    | 8.9         | 491    | 12    | BH521642           |
| 6          | 73.4  | 7.3         | 320    | 10    | L37476             |
| 7          | 68.8  | 6.9         | 749    | 10    | BE905170           |
| 8          | 68.2  | 6.8         | 595    | 12    | BH022954           |
| 9          | 66.6  | 6.7         | 593    | 10    | BG13327            |
| 10         | 65.2  | 6.5         | 868    | 10    | BE272517           |
| 11         | 63.8  | 6.4         | 679    | 9     | AU061011           |
| 12         | 63    | 6.3         | 375    | 10    | C68665             |
| 13         | 63    | 6.3         | 573    | 10    | BJ109423           |
| 14         | 63    | 6.3         | 637    | 10    | BJ116696           |
| 15         | 62.4  | 6.2         | 487    | 10    | BF704392           |
| 16         | 62.4  | 6.2         | 529    | 9     | AA896824           |
| 17         | 62.4  | 6.2         | 716    | 9     | BE055519 GA_Ea000  |

|    |      |     |      |    |          |                     |
|----|------|-----|------|----|----------|---------------------|
| 18 | 62.4 | 6.2 | 914  | 10 | BG917939 | BG917939 602820868  |
| 19 | 62.4 | 6.2 | 2738 | 11 | AK009815 | AK009815 Mus muscu  |
| 20 | 61.8 | 6.2 | 785  | 10 | BI920158 | BI920158 ESTP540093 |
| 21 | 60.8 | 6.1 | 681  | 10 | BE614098 | BE614098 601503838  |
| 22 | 60.6 | 6.1 | 243  | 10 | C94562   | C94562 C94562 Pig   |
| 23 | 60.2 | 6.0 | 466  | 10 | BG350414 | BG350414 091C01 Ma  |
| 24 | 59.8 | 6.0 | 863  | 10 | AT552716 | AT552716 mm09b08.y  |
| 25 | 58.8 | 5.9 | 566  | 10 | BI642658 | BI642658 T2S1024.T  |
| 26 | 58.4 | 5.8 | 706  | 9  | AI386274 | AI386274 mm68c05.y  |
| 27 | 57.2 | 5.7 | 887  | 12 | A2686939 | A2686939 ENTPN29TF  |
| 28 | 57.2 | 5.7 | 896  | 12 | BH150681 | BH150681 ENTPN77TR  |
| 29 | 57.2 | 5.7 | 927  | 12 | BH134897 | BH134897 ENTPN91TR  |
| 30 | 56.6 | 5.7 | 555  | 9  | AT452088 | AT452088 mm09b08.x  |
| 31 | 55.8 | 5.6 | 460  | 9  | AA919922 | AA919922 yf18g01.r  |
| 32 | 55.8 | 5.6 | 648  | 10 | BG762693 | BG762693 602734594  |
| 33 | 55.8 | 5.6 | 711  | 10 | BI693151 | BI693151 603344334  |
| 34 | 55.8 | 5.6 | 786  | 9  | AU002729 | AU002729 AUC02729   |
| 35 | 55.6 | 5.6 | 852  | 12 | A2530400 | A2530400 ENTPN11TF  |
| 36 | 55   | 5.5 | 495  | 9  | AT614714 | AT614714 v950c09.y  |
| 37 | 54.8 | 5.5 | 613  | 9  | AA444799 | AA444799 v950c09.y  |
| 38 | 54.4 | 5.4 | 616  | 9  | AU006028 | AU006028 AUC06028   |
| 39 | 53.6 | 5.4 | 360  | 10 | D36576   | D36576 CELK034E7F   |
| 40 | 53.6 | 5.4 | 1043 | 12 | CNS0145P | AL103735 Drosophila |
| 41 | 53.4 | 5.3 | 447  | 12 | A2924939 | A2924939 4906.1d55  |
| 42 | 52.8 | 5.3 | 985  | 12 | CNS07801 | AT433384 T7 end of  |
| 43 | 52.6 | 5.3 | 531  | 9  | AA786861 | AA786861 m6901a1.r  |
| 44 | 52.2 | 5.2 | 387  | 10 | Z14737   | Z14737 CEL2H12 Chr  |
| 45 | 52.2 | 5.2 | 928  | 12 | CNS00DKY | AL071865 Drosophila |

## ALIGNMENTS

RESULT 1  
AV939010

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

BASE COUNT

ORIGIN

Query Match

12.6% Score 125.8; DB 9; Length 678;

|                       |        |              |          |
|-----------------------|--------|--------------|----------|
| Best Local Similarity | 84.0%; | Pred. No.    | 3.9e-16; |
| Matches               | 142;   | Conservative | 0;       |
|                       |        | Mismatches   | 27;      |
|                       |        | Indels       | 0;       |
|                       |        | Gaps         | 0;       |

Qy 598 gcttgttccaggtaccatgttactaatltttgcaccaagtagcglttgyaactcca 6577  
||| ||||||||||||||||||||||||||||||||||||||||  
Db 495 GCGAGCTTGGGTACCATGTACTAATTTTTCGACCAAGTAGCGTTCGGAACGCCA 554

|    |  |     |  |      |
|----|--|-----|--|------|
| Oy |  | 658 | gaggaactaaatccttgcatacgacacaatyagtcytgtttccattcatgat   | 717  |
| Dd |  | 555 | GAGACTTAAATCCTTGATGCATAGACACAATGAGCGTTGGTTTCGTCTTAGCAT | 6144 |

QY 718 atgttcataagtaattagtccaathtaatttagctgttttactggtt 766  
||||| | | |||||  
Db 615 ATGTTCATAGTCAATTGCTCAAATAAATACCTTGACGGTTGAATGTT 663

RESULT 2

|            |   |        |      |        |                 |
|------------|---|--------|------|--------|-----------------|
| LOCUS      | BG240463  | 375 bp | mRNA | linear | EST 15-FEB-2001 |
| DEFINITION | OVI_30_D01.b1_A002 Ovary 1 (OVI) sorghum bicolor cDNA, mRNA |        |      |        |                 |

|           |            |             |
|-----------|------------|-------------|
| ACCESSION | BG240463   |             |
| VERSION   | BG240463.1 | GI:12775536 |
| KEYWORDS  | FST        |             |

ORGANISM      *Sorghum bicolor*  
Eukaryota: Viridiplantae

REFERENCE  
1 (bases 1 to 375)  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Sorghum.

**TITLE** An EST database from *Sorghum*: ovaries of varying immature stages  
**AUTHORS** L.H. Cordeonier-Pratt, M.-M. Gingle, A. Maisala, C. Sugman, M. and Pratt

unpublished (2000),  
Contact: Cordonnier-Pratt MM  
Department of Botany

THE UNIVERSITY OF GEORGIA  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860

Fax: 700 342 1803  
 Email: [mmpratt@uga.edu](mailto:mmpratt@uga.edu)  
 Sequences have been tr

below this quality is: the threshold for highest quality sequence is 20.  
Seq primer: JEN REV

Location/Qualifiers

/organism="Sorghum bicolor"  
/db\_xref="taxon:4558"

```
/note="Organ: Mix of ovaries of varying immature stages
from 8-week-old plants. Vector: pBluescript II from Lambda
```

from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN

|                       |        |                    |        |             |
|-----------------------|--------|--------------------|--------|-------------|
| Query Match           | 10.58; | Score 104.8;       | DB 10; | Length 375; |
| Best Local Similarity | 69.6%; | Pred. No. 8.9e-12; |        |             |
| Matched               | 143    | Genbank            | 1      | 0           |

Oy 271 aatacttagagatgcacatcgaatttagtgtgacctgcacaggaaccgaataat 330

Db 26 AATCATTCAGGATATATGAAACACATGTTGGATGAGTAGCCCGGAACCGAAGATAAAC 85

Db  
86 ACATATGCAACTTTAGSGATGAGTCTCCCAAGATATAAAAACTTGATTCATGCA 145

[illegible]

### RESULT 3

|            |   |        |      |        |                 |
|------------|---|--------|------|--------|-----------------|
| LOCUS      | AW065509  | 639 bp | mRNA | linear | EST 30-MAR-2000 |
| DEFINITION | 687002G09.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA |        |      |        |                 |

|             |   |            |
|-------------|---|------------|
| ACCESSION   | AW065909  | GI:6020981 |
| VERSION     | AW065909.1  |            |
| REVISIONS   | 1   |            |
| DATE        | 2009-01-21  |            |
| DESCRIPTION | Genomic DNA from <i>Escherichia coli</i> O157:H7 strain 26869, isolated from a bovine carcass in 1996. The genome was sequenced using a combination of Sanger and Illumina sequencing technologies. The genome is 4,859,800 bp in size and contains 11,000 genes. The genome is highly similar to other <i>E. coli</i> O157:H7 strains, including the 26869 strain. The genome is also highly similar to the 26869 strain, which is a reference strain for this serotype. The genome is highly similar to the 26869 strain, which is a reference strain for this serotype. The genome is highly similar to the 26869 strain, which is a reference strain for this serotype. |            |

| SOURCE ORGANISM        | STREPTONHUTINASE ACTIVITY |
|------------------------|---------------------------|
| <i>Zea mays</i> .      | 100                       |
| <i>Zea mays</i>        | 100                       |
| <i>Eukaryota</i> .     | 100                       |
| <i>Viridiniantae</i> . | 100                       |
| <i>Streptonhuta</i> .  | 100                       |
| <i>Tracheonhuta</i>    | 100                       |

REFERENCE  
1 (bases 1 to 639)  
spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.

**AUTHORS** waldbot, V.  
**TITLE** Maize ESTs from various cDNA libraries sequenced at Stanford University

JOURNAL  
unpublished (1999)  
COMMENT  
Contact: Walbot V  
Department of Biological Sciences

Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227

FAX: 030 723 0221  
 Email: walbot@stanford.edu  
 Plate: 687002 row: G column: 09.

```

1. .639
/organism="Zea mays

```

```

/db_xref="taxon:4577"
/clone_lib="687 - Early embryo from Delaware"

```

```

/dev_stage="14, 21, 28, and 35 days after pollination"
/lab_host="E. coli SOLR"

```

; Site\_2: EcoRI; Library was prepared by Stratagene using the Uni-ZAP XR system (Stratagene BN937328-12). Clones

(ampicillin resistance - use 100 micrograms/microliter).  
Developed from a pool of equal amounts of RNA from

pollination of the Illinois High Oil Maize Strain Cycle 90. This closed strain has been selected for high oil

| BASE COUNT | 1890s era open pollinated variety | Burr's White" |
|------------|-----------------------------------|---------------|
| 200 a      | 118 c                             | 151 g         |
|            |                                   | 170 t         |

Quincy Match 10 14, Score 100 0, PB 0, Toss 630.

Best Local Similarity 57.4%; Pred. No. 5.9e-11;  
Matches 201; Conservative 0; Mismatches 147; Indels 2; Gaps 1;

QY 127 atctgcgagcattccaacttcattgttctcgaggtgtgaactgcagcgatcttc 186

187 cattatgaagagatat--agttatcttltgtaacctacttgyaacttgaagtcctgag 244

Db 297 CAGGAGAAATACCATATGATGGGATTATTATTAAGATCCCTCGAAGAGGTAAGTATGTCT 356

\_\_\_\_\_

|            |   |   |                            |
|------------|---|---|----------------------------|
| Db         | 357   | TCAGCATTGCCGACACTTAAACGCACCAAAATCATCTTTGGATTATATGAAAACACATGTTCGGAA    | 416                        |
| Oy         | 305   | atcttgcacaggaaccgagaataaatcatatgatcgttaatttcaggatggagtgtttccaa        | 364                        |
| Db         | 417   | TGAGTAGGCCCGGACCCGAAGAATMAAACACATATGTAAACTTTAGGATGTAAGTCCTCCAA        | 476                        |
| Oy         | 365   | gaattaanaagcttggatcacaaatgcaatgacgcatatatgccaatccagagcattact          | 424                        |
| Db         | 477   | GAAATAAAAAACCCTTGATATCACATGCAATGCAATGCAATAATATGCAATCCAAAGCAGCACTATATT | 536                        |
| Oy         | 425   | atgcagacttggatggttatcacacaatccatctttttcttgataaacctt                   | 474                        |
| Db         | 537   | ATGGAAGCTTTTGATATACCATGTAACTATTTTTTTTCGCCCAAGTAGTGTCT                 | 586                        |
| RESULT     | 4   |   |                            |
| CNS00SBA   |   |   |                            |
| LOCUS      | CNS00SBA  | 386 bp  | DNA linear GSS 28-JUN-1999 |
| DEFINITION | Arabidopsis thaliana genome survey sequence T7 end of BAC T1H18 of TAU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence. |   |                            |
| ACCESSION  | AL088180  |   |                            |
| VERSION    | AL088180.1  | GI:5289320  |                            |
| KEYWORDS   | GSS.  |   |                            |
| SOURCE     | thale cress.  |   |                            |
| ORGANISM   | Arabidopsis thaliana  |   |                            |
|            | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  |   |                            |
|            | Spermatophytes; Magnoliophyta; eudicotyledons: core eudicots;   |   |                            |
|            | Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.   |   |                            |
| REFERENCE  | 1 (bases 1 to 386)  |   |                            |
| AUTHORS    | Salanoubat,M., Choinsne,N., Artiguenave,F., Brotlrier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.                             |   |                            |
| JOURNAL    | Unpublished   |   |                            |
| REFERENCE  | 2 (bases 1 to 386)  |   |                            |
| AUTHORS    | Genoscope.  |   |                            |
| TITLE      | Direct Submission   |   |                            |
| JOURNAL    | Submitted (25-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr                |   |                            |
|            | - Web : www.genoscope.cns.fr  |   |                            |
| FEATURES   | Location/Qualifiers   |   |                            |
| SOURCE     | 1..386  |   |                            |
|            | /organism="Arabidopsis thaliana"  |   |                            |
|            | /strain="Columbia"  |   |                            |
|            | /db_xref="taxon:3702"   |   |                            |
|            | /clone_lib="TAU"  |   |                            |
|            | /clone="T1H18"  |   |                            |
|            | /note="end : 77"  |   |                            |
| BASE COUNT | 107 a 71 c 78 g 130 t   |   |                            |
| ORIGIN     |   |   |                            |

|                       |                |                   |           |            |
|-----------------------|----------------|-------------------|-----------|------------|
| Query Match           | 9.2%           | Score 92.2        | DB 12     | Length 386 |
| Best Local Similarity | 57.5%          | Pred. No. 3,6e-09 |           |            |
| Matches 218           | Conservative 0 | Mismatches 138    | Indels 23 | Gaps 2     |

  

|    |     |   |     |
|----|-----|---|-----|
| QY | 379 | gatacaatgcagtgacaagataaaggcaatccaggagacattctactatgcaagcttggg    | 438 |
|    |     |   |     |
| Db | 1   | GGATATATATGCGCTTCATATTTGGCCATACAGAACATTCATATTTATGCCAGCTTTGGG    | 60  |
|    |     |   |     |
| QY | 439 | tattcacacaatcacatlttttttcgtataacactcttcaccatcttggagctatcaccc    | 498 |
|    |     |   |     |
| Db | 61  | TATTTCTTTTAA-CCTTAGGAGCCTCTAGAGAACCTATATTTGGCATTTATATTTCT       | 119 |
|    |     |   |     |
| QY | 499 | taatcttcacagcacataaaatatcttgataataacccattatagatatatagtaacact    | 558 |
|    |     |   |     |
| Db | 120 | TGCACAACTATATGCGTTTATGGCTTTTGGACCATGATGCTACTGTATCTTCAAGCAACT    | 179 |
|    |     |   |     |
| QY | 559 | aacacttagtattctgaaaaagaatcatctttatgct-----ttgct                 | 596 |
|    |     |   |     |
| Db | 180 | AGTACTATATGATATCTGCATTTCTTATTTGGTTTTCTTACTAGGAACAATGCCCATTTTCTA | 239 |
|    |     |   |     |
| QY | 597 | ggactgttccaggtaccacatgcttaactaatlttttgcacaaagtagccglttggaaacct  | 656 |
|    |     |   |     |

|            |  |  |                            |
|------------|--|--|----------------------------|
| Db         | 240  | TTATGAATGSCAGCTACACATGTCACCAACTTTTTCGCCCAAGCAATCGCTGCGGACCCC | 299                        |
| Qy         | 657  | agagagactaaatccctgatacgatagacagacatgagcttggttgctgttctatgga   | 716                        |
|            |  |  |                            |
| Db         | 300  | AGAGAGACTAAATACCTGATGATAGAGAGCTCAGACAGTTAGCGCTGGTGTCTGATGGA  | 359                        |
| Qy         | 717  | tattgttcataagtaatta  | 735                        |
|            |  |  |                            |
| Db         | 360  | TATCGTTTCATAGTAGTAGTA  | 378                        |
| RESULT     | 5  |  |                            |
| LOCUS      | BH521642   | 491 bp   | DNA linear GSS 13-DEC-2001 |
| DEFINITION | BOGTD27TR BOGT Brassica oleracea genomic clone BOGTD27, DNA  |  |                            |
| ACCESSION  | BH521642   |  |                            |
| VERSION    | BH521642.1   | GI:17729727  |                            |
| KEYWORDS   | GSS.   |  |                            |
| SOURCE     | Brassica oleracea.   |  |                            |
| ORGANISM   | Brassica oleracea  |  |                            |
| REFERENCE  | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;<br>Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;<br>Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica. |  |                            |
| AUTHORS    | 1 (bases 1 to 491)   |  |                            |
| TITLE      | Tom,C.D., Van Aken,S., Uteerback,T. and Fraser,C.M.  |  |                            |
| JOURNAL    | Whole genome shotgun sequencing of Brassica oleracea<br>unpublished (2001)   |  |                            |
| COMMENT    | Other_GSSs: BOGTD27TF<br>Contact: Chris Town   |  |                            |
|            | TIGR   |  |                            |
|            | 9712 Medical Center Drive, Rockville, MD 20850, USA.   |  |                            |
|            | Tel: 301-838-3523  |  |                            |
|            | Fax: 301-838-0208  |  |                            |
|            | Email: cdtown@tigr.org   |  |                            |
|            | DNA is from a doubled haploid provided by Tom Osborn.  |  |                            |
|            | Seq primer: TR   |  |                            |
|            | Class: sheared ends.   |  |                            |
| FEATURES   | Location/Qualifiers  |  |                            |
| source     | 1..491   |  |                            |
|            | /organism="Brassica oleracea"  |  |                            |
|            | /strain="TO1000DH3"  |  |                            |
|            | /db_xref="taxon:3712"  |  |                            |
|            | /clone="BOGTD27"   |  |                            |
|            | /clone.lib="BOGT"  |  |                            |
|            | /note="Vector: pROS1, Site_1: BstXI; 2-3 kb sheared<br>genomic DNA inserted into pROS1 using BstXI linkers"  |  |                            |
| BASE COUNT | 144 a 89 c 99 g 159 t  |  |                            |
| ORIGIN     |  |  |                            |

| Query Match           | 8.9%   | Score 89          | DB 12    | Length 491        |
|-----------------------|--|-------------------|----------|-------------------|
| Best Local Similarity | 80.6%  | Fred. No. 1.6e-08 |          |                   |
| Matches 104           | Conservative 0   | Mismatches 25     | Indels 0 | Gaps 0            |
| Qy 312                | acaggaacccaagataaattcatatgctcaattttagggatgagtggttgccaagaattaa    | 371               |          |                   |
|                       |  |                   |          |                   |
| Db 294                | ACAGGACCAATGATTAATATCATATGCCCACCTTTAGAGATGATGTACTCCCCGTATAAA     | 353               |          |                   |
| Qy 372                | aagcttggatacaatgacagtgcaagaataatgccaatccagagcatctactatgcaag      | 431               |          |                   |
|                       |  |                   |          |                   |
| Db 354                | AAAGCTAGAGCTTAAATAGCTGTTCAGATATATGSCCATTCAGAGACATTCCTACTATGACCAG | 413               |          |                   |
| Qy 432                | ctttgggta 440  |                   |          |                   |
|                       |  |                   |          |                   |
| Db 414                | CTTTGGGTA 422  |                   |          |                   |
| RESULT 6              |  |                   |          |                   |
| LOCUS 137476          |  |                   |          |                   |
| DEFINITION L37476     |  | 320 bp            | mRNA     | linear            |
|                       | BNAESTTQ Mustard flower buds                                     |                   |          |                   |
|                       | Brassica rapa  |                   | cdna     | clone F0956, mRNA |
|                       | sequence.  |                   |          |                   |

|                       |  |
|-----------------------|--|
| ACCESSION             | L37476   |
| VERSION               | L37476.1   |
| KEYWORDS              | GI:887054  |
| SOURCE                | EST.   |
| ORGANISM              | Brassica rapa.<br>Brassica rapa<br>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;<br>Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;<br>Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.            |
| REFERENCE             | 1 (bases 1 to 320)<br>Lim,C.O., Kim,H.Y., Kim,M.G., Lee,S.I., Chung,W.S., Park,S.H.,<br>Hwang,I. and Cho,M.J.<br>Expressed sequence tags of Chinese cabbage flower bud cDNA<br>Plant Physiol. 111 (2), 577-588 (1996)                          |
| TITLE                 | Contact: Lim,C.O., Kim,M.G., Hwang,I. and Cho,M.J.   |
| JOURNAL               | Plant Molecular Biology and Biotechnology Research Center  |
| MEDLINE               | Email: pmbrc@nongae.gsnu.ac.kr.  |
| COMMENT               | Location/Qualifiers  |
| FEATURES              | 1..320   |
| SOURCE                | /organism="Brassica rapa"<br>/strain="pekinensis"<br>/db_xref="taxon:3711"<br>/clone="F0956"<br>/clone_lib="Mustard flower buds"<br>/note="Devel_stage = flower bud "  |
| BASE COUNT            | 77 a 72 c 79 g 92 t  |
| ORIGIN                |  |
| Query Match           | 7.3%; Score 73.4; DB 10; Length 320;   |
| Best Local Similarity | 72.5%; Pred. No. 2.8e+05;  |
| Matches               | 95; Conservative 0; Mismatches 36; Indels 0; Gaps 0;   |
| OY                    | 598 gccgtgccaggtaccacggtactaatcttttgcaccagaftgcgttttgaaactca 657<br>   |
| Db                    | 6 gctaccttgggtatcancagtataccaacattttccctccctacagcgccgttggaacct 65<br>  |
| OY                    | 658 gaggaactaaaactctgatcgatgagacatgagcttgtgttcgtttcttatgat 717<br>   |
| Db                    | 66 AATGACCCTTAATTCCTTAAATAGCAAGCGCTCATGTAGCTGGCTGCTCCTAATGAT 125<br>   |
| OY                    | 718 attgttcatat 728<br>  |
| Db                    | 126 ATTGTGCACAG 136  |
| RESULT                | 7  |
| BE905170              | mRNA linear EST 20-OCT-2000  |
| LOCUS                 | BE905170 749 bp mRNA clone IMAGE:3901214 5',   |
| DEFINITION            | 601499362P1 NIH_MGC-70 Homo sapiens CDNA clone IMAGE:3901214 5',<br>mRNA sequence.   |
| ACCESSION             | BE905170   |
| VERSION               | BE905170.1   |
| KEYWORDS              | GI:10398185  |
| SOURCE                | EST.   |
| ORGANISM              | human.<br>Homo sapiens<br>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;<br>Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.   |
| REFERENCE             | 1 (bases 1 to 749)<br>NIH-MGC http://mgc.ncl.nih.gov/<br>National Institutes of Health, Mammalian Gene Collection (MGC)<br>Unpublished (1999)<br>Contact: Robert Strausberg, Ph.D.<br>Email: cgabbs-remail.nih.gov<br>Tissue Procurement: ATCC |
| TITLE                 | CDNA Library Preparation: Life Technologies, Inc.<br>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  |
| JOURNAL               | DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be<br>found through the I.M.A.G.E. Consortium/LNL at:<br>http://image.lnl.gov   |
| COMMENT               | Plate: LHAM9701 row: 1 column: 15<br>High quality sequence stop: 691.  |

| FEATURES              | SOURCE  | Location/Qualifiers  |
|-----------------------|---|--|
|                       |   | 1..749   |
|                       |   | /organism="Homo sapiens"   |
|                       |   | /db_xref="taxon:9606"  |
|                       |   | /clone="IMAGE:3901214"   |
|                       |   | /clone_1id="NH.MGC.70"   |
|                       |   | /tissue_type="epithelioid carcinoma"   |
|                       |   | /lab_host="DH10B (phage-resistant)"  |
|                       |   | /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.1 kb. Library constructed by life technologies." |
| BASE COUNT            | 225 a   | 137 c 165 g 222 t  |
| ORIGIN                |   |  |
| Query Match           | 6.9%;   | Score 68.8; DB 10; Length 749;   |
| Best Local Similarity | 63.1%;  | Pred. No. 0.00024;   |
| Matches 106;          | Conservative 0;   | Mismatches 62; Indels 0; Gaps 0;   |
| QY 309                | tgcacaggaacccaagaataataatcatatgctaatcttgtaggaatgaggtgtgccagaat 368    |  |
| Db 337                | TTCCCATGAAAGAAAGAGACGCTTCTTATAAACATTTTACATGCAATGTACTACCAAGAT 386      |  |
| QY 369                | taaaagcttgatatacaatctgcagtcagacagataatgcaatccagagacataactatgc 428     |  |
| Db 387                | CAAAAGCCCTTGGAATCAACAGCTCATCTTCACTTATGCAATGCAATGAGCAATGCTTACTATGC 446 |  |
| QY 429                | aagcttggtatcacacaatccatctctttctctgtacacactctca 476                    |  |
| Db 447                | CAGCTTGTGTTACCAATACACAGCTTCTTTTGCAAGCTTCACGCCCTTA 494                 |  |
| RESULT 8              | BH022954  | 595 bp DNA linear GSS 29-JUN-2001  |
| LOCUS                 | BH022954/c  |  |
| DEFINITION            | GH_MBB00030322r Gossypium hirsutum L. Gossypium hirsutum genomic      |  |
| ACCESSION             | BH022954  |  |
| VERSION               | BH022954.1  | GI:14577242  |
| KEYWORDS              | GSS.  |  |
| SOURCE                | upland cotton.  |  |
| ORGANISM              | Gossypium hirsutum  |  |
|                       | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;    |  |
|                       | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;          |  |
|                       | Rosidae; eucosids II; Malvales; Malvaceae; Gossypium.                 |  |
| REFERENCE             | 1 (bases 1 to 595)  |  |
| AUTHORS               | Tomkins,J.P., Peterson,D.G., Yang,T.J., Main,D., Wilkins,T.A.,        |  |
|                       | Paterson,A.H., and Wing,R.A.  |  |
| TITLE                 | Development of Genomic Resources for Cotton (Gossypium hirsutum       |  |
|                       | L.): BAC Library Construction, Preliminary STC Analysis, and          |  |
|                       | Identification of Clones Associated With Fiber Development            |  |
| JOURNAL               | Unpublished (2001)  |  |
| COMMENT               | Contact: Wing RA  |  |
|                       | Clemson University Genomics Institute                                 |  |
|                       | Clemson University  |  |
|                       | 100 Jordan Hall, Clemson, SC 29634, USA                               |  |
|                       | Tel: 864 656 7288   |  |
|                       | Fax: 864 656 4293   |  |
|                       | Email: twing@clemson.edu  |  |
|                       | Class: BAC ends   |  |
|                       | High quality sequence stop: 569.                                      |  |
| FEATURES              | Location/Qualifiers   |  |
| SOURCE                | 1..595  |  |
|                       | /organism="Gossypium hirsutum"  |  |
|                       | /cultivar="Maxxa"   |  |
|                       | /db_xref="taxon:3635"   |  |
|                       | /clone="GH_MBB00030322r"  |  |
|                       | /clone_1id="Gossypium hirsutum L."                                    |  |
|                       | /tissue_type="Young leaves"   |  |
|                       | /lab_host="E. coli"   |  |
|                       | /note="Vector: pCUGIBAC-1; Site_1: HindIII; Site_2: NotI;             |  |
|                       | For more details on library preparation, ordering clones              |  |

cgatagagcacatgagcttggttgccttgcttcttat

**KEYWORDS**

|                       |  |
|-----------------------|--|
| SOURCE                | Dictyostelium discoideum.  |
| ORGANISM              | Dictyostelium discoideum   |
| REFERENCE             | Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.   |
| AUTHORS               | Morioto, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Maeda, B. N., Pl, M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Mitsuda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.  |
| TITLE                 | Developmental cDNA in Dictyostelium discoideum   |
| JOURNAL               | Unpublished (1998)   |
| COMMENT               | Contact: Hideko Urushihara<br>Institute of Biological Sciences<br>University of Tsukuba<br>3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan<br>Email: d402huesakura.cc.tsukuba.ac.jp<br>PROJECT = Dictyostelium discoideum cDNA project in Japan.<br><br>Location/Qualifiers<br>1..679 |
| FEATURES              | /organism="Dictyostelium discoideum"<br>/strain="AX4"<br>/db_xref="taxon:44689"<br>/clone="SLC606"<br>/clone_id="Dictyostelium discoideum SL (H.Urushihara)"<br>/dev_stage="slug"  |
| BASE COUNT            | 202 a 107 c 162 g 208 t  |
| ORIGIN                |  |
| Query Match           | 6.4%; Score 63.8; DB 9; Length 679;  |
| Best Local Similarity | 67.9%; Pred. No. 0.0026;   |
| Matches               | 89; Conservative 0; Mismatches 42; Indels 0; Gaps 0;   |
| OY                    | 598 gctgttcacagtgaccatgtaacttatttttcaccacaagtagccgtttggaaacctca 657  |
| Db                    | 248 GCATCGTTGGTTCATCAAGTGACAAATTCTTTCGATCACAGAGAATGTGGCACACCG 307  |
| OY                    | 658 gaggaactaaatacctctgatcgatagaacacatgagcttgttgccttctaagtat 717   |
| Db                    | 308 GAGGAGGTGAAGACATATTGATTAAGACACGAGATGGAGATGTGTGTGTTTTAAGAC 367  |
| OY                    | 718 attgtcatcat 728  |
| Db                    | 368 GTGTGCATAG 378   |
| RESULT 12             |  |
| LOCUS                 | C68665 375 bp mRNA linear EST 23-SEP-1997  |
| DEFINITION            | C68665 Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA clone yk306c3 5', mRNA sequence.   |
| ACCESSION             | C68665   |
| VERSION               | C68665.1 GI:2430021  |
| KEYWORDS              | EST.   |
| SOURCE                | Caenorhabditis elegans.  |
| ORGANISM              | Caenorhabditis elegans.<br>Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitae; Rhabditoidea ; Rhabdilitidae; Peloderinae; Caenorhabditis.  |
| REFERENCE             | 1 (bases 1 to 375)<br>Kohara, Y., Motobashi, T., Tabara, H., Watanabe, H., Sugimoto, A., Sano , M., Miyata, A. and Nishigaki, A.   |
| AUTHORS               | Expression map of the C.elegans genome   |
| TITLE                 | Unpublished (1996)   |
| JOURNAL               | Contact: Yuji Kohara   |
| COMMENT               | Genome Biology Lab.<br>National Institute of Genetics<br>Yata 1111, Mishima, Shizuoka 411, Japan<br>Tel: 81-559-81-6854<br>Fax: 81-559-81-6855<br>Email: yokohara@lab.nig.ac.jp.<br><br>Location/Qualifiers<br>1..375  |
| FEATURES              | /organism="Caenorhabditis elegans"<br>/strain="CB1489 him-8(e1489)"<br>/db_xref="taxon:6239"   |

|                       |  |                        |                   |                        |
|-----------------------|--|------------------------|-------------------|------------------------|
|                       | Query Match  | 6.3%; Score 63; DB 10; | Length 375;       |                        |
|                       | Best Local Similarity  | 63.6%;                 | Pred. No. 0.0039; |                        |
| Matches               | 96; Conservative   | 0; Mismatches          | 55; Indels        | 0; Gaps<br>0;          |
| OY                    | 322 aagataaatcatatgccaattttagtgagtgtttgccaaagaattaaggctggga  | 381                    |                   |                        |
| Dd                    | 76 AAAATTAAACATACTGTAATCGGAATTCGGTCAAGCATGTCTTGCCGAGAATAACAACCAAGGA  | 135                    |                   |                        |
| Oy                    | 382 tacatatgcacgcagcagatatgatccactcaggcaagtcattcatactaqtgcaaagcttgggtat  | 441                    |                   |                        |
| Dd                    | 136 TACAAATGCATAATTCACGTTATGTGGCAGTAGTGGACGACTGTTATTATGCAATCTTTGGATAT  | 195                    |                   |                        |
| Oy                    | 442 tcacacaatcacattttctctgtctaacacc  | 472                    |                   |                        |
| Dd                    | 196 CAGCTTTGGAAGCTTTTTCGCTGTTCAGTCC  | 226                    |                   |                        |
| RESULT 13             | Bj109423   | 573 bp                 | mRNA              | linear EST 23-JAN-2002 |
| LOCUS                 | Bj109423 unpublished oligo-capped cDNA library, C. elegans l1 stage  |                        |                   |                        |
| DEFINITION            | Caeonorbadditis elegans CDNA clone YK1119C01 5' mRNA sequence.   |                        |                   |                        |
| ACCSSION              | Bj109423   |                        |                   |                        |
| VERSION               | Bj109423..1 GI:18269452  |                        |                   |                        |
| KEYWORDS              | EST,   |                        |                   |                        |
| SOURCE                | Caeonorbadditis elegans.   |                        |                   |                        |
| ORGANISM              | Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabdioidea   |                        |                   |                        |
| AUTHORS               | Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.<br>and Sugano,S.   |                        |                   |                        |
| REFERENCE             | A complementary view of the C.elegans genome   |                        |                   |                        |
| JOURNAL               | Unpublished (2002)   |                        |                   |                        |
| COMMENT               | Contact: Tadasu Shin-i<br>Center For Genetic Resource Information<br>National Institute of Genetics<br>111 Yata, Mishima, Shizuoka 411-8540, Japan<br>Tel: 81-559-81-6856<br>Fax: 81-559-81-6855<br>Email: tshini@genes.nig.ac.jp. |                        |                   |                        |
| FEATURES              | Location/Qualifiers  |                        |                   |                        |
| source                | 1..573   |                        |                   |                        |
| /organism=            | "Caenorhabditis elegans"   |                        |                   |                        |
| /strain=              | "N2"   |                        |                   |                        |
| /db_xref=             | "taxon:6239"   |                        |                   |                        |
| /clone_1id=           | "YK1119C01"  |                        |                   |                        |
| /clone_1lib=          | "unpublished oligo-capped cdna library, C. elegans l1 stage"   |                        |                   |                        |
| /sex=                 | "hermaphrodite"  |                        |                   |                        |
| /tissue-type=         | "whole animal"   |                        |                   |                        |
| /dev_stage=           | "L1"   |                        |                   |                        |
| BASE COUNT            | 169 a      121 c      127 g      155 t      1 others   |                        |                   |                        |
| ORIGIN                |  |                        |                   |                        |
| Query Match           | 6.3%; Score 63; DB 10;   | Length 573;            |                   |                        |
| Best Local Similarity | 63.6%;   | Pred. No. 0.0039;      |                   |                        |
| Matches               | 96; Conservative   | 0; Mismatches          | 55; Indels        | 0; Gaps<br>0;          |
| OY                    | 322 aagataaatcatatgccaattttagtgagtgtttgccaaagaattaaggctggga  | 381                    |                   |                        |
| Dd                    | 309 AAATTAACACATATCGTAATTCGCTGACGATGTCTTGGCCGAGATACAAAACCAAGGA   | 368                    |                   |                        |

|                       |   |   |                             |
|-----------------------|---|---|-----------------------------|
| Oy                    | 382   | tacaatgcagtcagataatgcaatccgaacggcattcaactatgcagaacttggatat    | 441                         |
| Dd                    | 369   | TGACATGCAGATTCACTTGATGGCAGTGAATGACGACCGTTTATTATGATCATTTTGATAT | 428                         |
| Oy                    | 442   | tcacacaatccaattttcttgtatatacac                                | 472                         |
| Dd                    | 429   | CAGGTTCGAACTTTTTCGCTGTTCCAGTC                                 | 459                         |
| RESULT                | 14  |   |                             |
| Bull6696              |   |   |                             |
| LOCUS                 | Bull6696  | 637 bp  | mRNA linear EST 23-JAN-2002 |
| DEFINITION            | Bull6696 unpublished oligo-capped cDNA library; C. elegans L1 stage   |   |                             |
| ACCESSION             | Caenorhabditis elegans cDNA clone yk1206c05 5', mRNA sequence.  |   |                             |
| VERSION               | Bull6696  |   |                             |
| KEYWORDS              | Bull6696.1 GI:18276808  |   |                             |
| SOURCE                | EST.  |   |                             |
| ORGANISM              | Caenorhabditis elegans.<br>Caenorhabditis elegans.<br>Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea<br>; Rhabditidae; Pelodierinae; Caenorhabditis.<br>1 (bases 1 to 637)<br>Kohara,Y., ShIn-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.<br>and Sugano,S.<br>A complementary view of the C.elegans genome<br>Unpublished (2002)<br>Contact: Tadasu ShIn-i<br>Center For Genetic Resource Information<br>National Institute Of Genetics<br>1111 Yata, Mishima, Shizuoka 411-8540, Japan<br>Tel: 81-559-81-6856<br>Fax: 81-559-81-6855<br>Email: tshin@genes.nig.ac.jp.<br>Location/Qualifiers<br>1..637<br>/organism="Caenorhabditis elegans"<br>/strain="N2"<br>/db_xref="taxon:6239"<br>/clone=yk1206c05<br>/clone_lib="unpublished oligo-capped cDNA library, C.<br>elegans L1 stage"<br>/sex="hermaphrodite"<br>/tissue-type="whole animal"<br>/dev_stage="L1" |   |                             |
| FEATURES              | source  |   |                             |
| BASE COUNT            | 190 a 130 c 142 g 174 t   | 1 others  |                             |
| ORIGIN                |   |   |                             |
| Query Match           | 6.3%;   | Score 63;   | DB 10; Length 637;          |
| Best Local Similarity | 63.6%;  | Pidmat No. 0.0039;  |                             |
| Matches               | 96; Conservative 0;   | Mismatches 55;  | Indels 0; Gaps 0;           |
| Oy                    | 322   | aagataaatctcataagcgaatttgaggatgggtgcccgaagattaagaagcttga      | 381                         |
| Dd                    | 345   | AAAATTAAACACTATTCGTAATTCGCTGACGATGCTTTCGCGAGAATACAAAGAAGGA    | 404                         |
| Oy                    | 382   | tacaatgcagtcagataatgcaatccgaacggcattcaactatgcagaacttggatat    | 441                         |
| Dd                    | 405   | TGACATGCAGATTCACTTGATGGCAGTGAATGACGACCGTTTATTATGATCATTTTGATAT | 464                         |
| Oy                    | 442   | tcacacaatccaattttcttgtatatacac                                | 472                         |
| Dd                    | 465   | CAGGTTCGAACTTTTTCGCTGTTCCAGTC                                 | 495                         |
| RESULT                | 15  |   |                             |
| Bull704392            |   |   |                             |
| LOCUS                 | Bull704392  | 487 bp  | mRNA linear EST 22-DEC-2000 |
| DEFINITION            | MI-P-E6-acc-d-08-1-UM.s1 MI-P-E6 Sus scrofa cDNA clone  |   |                             |
| ACCESSION             | Bull704392  |   |                             |
| VERSION               | Bull704392.1  | GI:11989800   |                             |
| KEYWORDS              | EST.  |   |                             |

[illegible]





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 07:29:49 ; Search time 3917.59 Seconds  
(without alignments)  
5347.028 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_9500\_10500

Perfect score: 1001

Sequence: 1 tccgacatagacacatga.....ctggacatagctttttg 1001

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_com: \*  
5: gb\_ov: \*  
6: gb\_pal: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_cm: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_inv: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Query Match | Length DB | ID | Description |
|------------|-------------|-----------|----|-------------|
|------------|-------------|-----------|----|-------------|

|    |       |       |        |    |           |          |
|----|-------|-------|--------|----|-----------|----------|
| 1  | 1001  | 100.0 | 11463  | 6  | AX031271  | Sequence |
| 2  | 978   | 97.7  | 11475  | 8  | AF338431  | Aegilops |
| 3  | 223.2 | 22.3  | 23449  | 8  | AF072725  | Azorella |
| 4  | 136.4 | 13.6  | 2726   | 8  | AF338432  | Triticum |
| 5  | 136.4 | 13.6  | 2970   | 8  | AF286319  | Triticum |
| 6  | 136.4 | 13.6  | 2970   | 8  | TASBA2    | Triticum |
| 7  | 134.8 | 13.5  | 2853   | 8  | TA066376  | Triticum |
| 8  | 130   | 13.0  | 2554   | 8  | AF064560  | Triticum |
| 9  | 122   | 12.2  | 2795   | 8  | ZMG65948  | Triticum |
| 10 | 118.8 | 11.9  | 3015   | 6  | E14723    | Rice     |
| 11 | 118.8 | 11.9  | 3015   | 6  | E14723    | Rice     |
| 12 | 118.8 | 11.9  | 3015   | 6  | AB023498  | Oryza    |
| 13 | 110.8 | 11.1  | 2725   | 6  | AR106495  | Oryza    |
| 14 | 110.8 | 11.1  | 2725   | 6  | AR106495  | Oryza    |
| 15 | 107.8 | 10.8  | 419    | 8  | MZGLUCTRN | Oryza    |
| 16 | 106   | 10.6  | 2918   | 8  | AF51191   | Oryza    |
| 17 | 106   | 10.6  | 2918   | 8  | RI08C83   | Oryza    |
| 18 | 98    | 9.8   | 2819   | 6  | E08183    | Oryza    |
| 19 | 95.2  | 9.5   | 2913   | 6  | AF064561  | Oryza    |
| 20 | 94.8  | 9.5   | 3360   | 8  | AB029548  | Oryza    |
| 21 | 94.2  | 9.4   | 3074   | 6  | AR123355  | Oryza    |
| 22 | 93.6  | 9.4   | 2982   | 8  | STU011888 | Solanum  |
| 23 | 93.6  | 9.4   | 3003   | 6  | A58162    | Solanum  |
| 24 | 93.6  | 9.4   | 3090   | 6  | A92164    | Solanum  |
| 25 | 92.8  | 9.3   | 2529   | 6  | A58167    | Solanum  |
| 26 | 92.6  | 9.3   | 1393   | 6  | AR123356  | Solanum  |
| 27 | 92    | 9.2   | 2493   | 6  | STSBET1   | Solanum  |
| 28 | 92    | 9.2   | 2523   | 8  | STU011889 | Solanum  |
| 29 | 92    | 9.2   | 2955   | 8  | STU011885 | Solanum  |
| 30 | 92    | 9.2   | 2992   | 8  | STU011890 | Solanum  |
| 31 | 92    | 9.2   | 3033   | 6  | A58164    | Solanum  |
| 32 | 91.6  | 9.2   | 3549   | 6  | PSSBERGEN | Solanum  |
| 33 | 90.4  | 9.0   | 2563   | 6  | AX256072  | Solanum  |
| 34 | 90.4  | 9.0   | 2576   | 6  | A58166    | Solanum  |
| 35 | 90.4  | 9.0   | 2578   | 6  | A58169    | Solanum  |
| 36 | 90.4  | 9.0   | 2975   | 6  | A58163    | Solanum  |
| 37 | 90.4  | 9.0   | 3331   | 6  | A58168    | Solanum  |
| 38 | 88.6  | 8.9   | 102897 | 8  | ATF17C15  | Solanum  |
| 39 | 88.4  | 8.8   | 2517   | 8  | AB042937  | Solanum  |
| 40 | 88.4  | 8.8   | 3123   | 8  | AB071286  | Solanum  |
| 41 | 85.2  | 8.5   | 2542   | 8  | AT022428  | Solanum  |
| 42 | 81.6  | 8.2   | 7020   | 8  | ATSBET2   | Solanum  |
| 43 | 81.6  | 8.2   | 101365 | 8  | AC006919  | Solanum  |
| 44 | 80.8  | 8.1   | 2668   | 8  | ATU18817  | Solanum  |
| 45 | 67.4  | 6.7   | 2872   | 10 | BC017541  | Mus      |

#### ALIGNMENTS

RESULT 1  
AX031271  
LOCUS AX031271 11463 bp DNA linear PAT 20-SEP-2000  
DEFINITION Sequence 10 from Patent WO9914314.  
ACCESSION AX031271  
VERSION AX031271.1 GI:10278603  
KEYWORDS  
SOURCE  
ORGANISM  
Aegilops tauschii.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Pooidae; Triticaceae; Aegilops.

REFERENCE  
1 (bases 1 to 11463)  
Li, Z., Morell, M. and Rahman, S.  
Regulation of gene expression in plants  
Patent: WO 9914314-A 10 25-MAR-1999;  
GOODMAN FIELDER LTD (AU) ; LI ZHONGYI (AU) ; MORELL MATTHEW (AU) ;  
RAHMAN SADEOUR (AU) ; UNIV AUSTRALIAN (AU) ; COMMW SCIENC INT RES

FEATURES  
source  
1. 11463  
Location/Qualifiers  
/organism="Aegilops tauschii"  
/db\_xref="taxon:37682"

misc\_feature 1. 11463  
 /product="COMPLETE SEQUENCE OF THE STARCH BRANCHING ENZYME II GENE"  
 BASE COUNT 3093 a 2253 c 2422 g 3689 t  
 ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 11463;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-208;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tccgacatagacacgcatgacacatatacaagctctctcttctgacagatatagtatattca 60  
 DB 9500 TCCGACATAGAACACGACATGACATATCAAGCTCTCTTGTGCAGAGATATGATGATTCA 9559  
 QY 61 tgcgtctggtagagcttcaactcttcgcatgcatcgtgtgcatagatattcaataatgat 120  
 DB 9560 TGGCTCTGGATAGGCTTCAACTCTTCGCAATTCGATCGTGGCATTCACATTAATAATGAT 9619  
 QY 121 caggcttgcacacatgggtttaggtgtgtaagagctatcttaacttcaatggaatgaatt 180  
 DB 9620 CAGGCTTGCACCAATGGGTTTAGGGGTGAGAGCGTATCTTAATCTCATGGGAAATGAGATT 9679  
 QY 181 tgggacatccggtgcaagctcttaacaacattatgcattctgcattgattgattactgta 240  
 DB 9680 TGGGATCTCTGTCAGTCTTTACCAACATTATGCAATTCGATCGATGATGATTTACTGTA 9739  
 QY 241 attggaacacagctcttcttccacatgcatgcatgcatgcatgcatgcatgcatgcatg 300  
 DB 9740 ATTTGACCAACAGCTTTCTTTTTCACATTCATGATATGATATGATATGATATGATATG 9799  
 QY 301 gaagcttaactcttacttctgacagaaatgtagatattccaaagagcccaacaactct 360  
 DB 9800 GAAGTAACTCTCTAATTTACTTGGCAGATGATGATTTTCCAGAGCCCAACAACACTCT 9859  
 QY 361 tccaacgagcaagctctcccccgtggaataacaatagttatgataaatagcgcgtagat 420  
 DB 9860 TCCAACCGCAAGTCTCCCTCGCAAAATTAACATGATGATGATGATGATGATGATGATG 9919  
 QY 421 ttgactcttgtaagctttagctgtgcttactatattccactactagattctttaggcatt 480  
 DB 9920 TTGATCTTGTAGATTTAGCTGTGCTATTAATTCCTCACTAGATGATCTTTTATGGCCATT 9979  
 QY 481 tattcttgatgaaatcaataatgctttagtaagaaagcaaatgcttctttagattttg 540  
 DB 9980 TATTTCTGATGAATCATATATGTTTGTAGAAAGATCAACATTCCTTTGTAAGTTTG 10039  
 QY 541 tagagcttaacaataagtagtgtgtgtagaggtgtgtgtagatcataaataatcatgattttt 600  
 DB 10040 TAGACGTTTAACTAGTATGTTGTGAGAGTGTGATCATTAATAAATTCATGATTTT 10099  
 QY 601 gcaagggaatcacagatttcttagatatatcgtgtagatgcaagatgcatcaggcaatga 660  
 DB 10100 GCAGGAGATCAGATTTTCTTATGATATCGTGTGATGCAAGATTCGATCAGGCAATGCA 10159  
 QY 661 gcatcttgaggaataataatgaggtatgtagtctgttcttcttcttcttcttcttcttct 720  
 DB 10160 GCATCTTGAGGAAAAAATATGCGGTATGTCACCTGGTTCTCTTGTGATCAACAGTAC 10219  
 QY 721 agtttaagctgaagctcttcaagtggtataaaaaagctgtagaataatctctgttaata 780  
 DB 10220 AGTTTAAGCTGAGTCTCTTCAAGTGTGAAAAAAGTGTGATTAATTAATCTGTAATGAGA 10279  
 QY 781 tgaataagctgcaaaagcgagagctggaattgcttcttccaaactatcttcttaagtc 840  
 DB 10280 TGAATAAGCTGCAAAAGGCGAGCTGGAATTCCTTTTCAACCAAACTATTTCTTAAGTGC 10339  
 QY 841 ttggtatgtagatataacagacacgacacgacacgacacgacacgacacgacacgacacg 900  
 DB 10340 TTGGTATGATATACATATACAGACACGACATGATATGATGATGATGATGATGATGATG 10399  
 QY 901 ccagatgcttccaggaataacatgagaagataaggtgtagatcatctctcaaaagagagattt 960

DB 10400 CCAGATGTTTCACGGAAGCAATGAGAGATAGATGATCATCTCCCAAAAGAGAGATT 10459  
 QY 961 ggtattgcttccactccactgagacatagctttttg 1001  
 DB 10460 GGTATTGTTTCAACTCTGAGCAATGATCTTTTGT 10500

RESULT 2  
 AF338431 11475 bp DNA linear PLN 27-MAR-2001  
 LOCUS  
 DEFINITION Aegilops tauschii starch branching enzyme Iia gene, complete cds.  
 ACCESSION AF338431  
 VERSION AF338431.1 GI:13447949  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Aegilops tauschii.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Pooideae; Triticeae; Aegilops.  
 1 (bases 1 to 11475)  
 Rahman, S., Regina, A., Li, Z., Mukai, Y., Yamamoto, M.,  
 Kosar, Hashemi, B., Abrahams, S. and Morell, M. K.  
 Comparison of Starch-Branching Enzyme Genes Reveals Evolutionary  
 Relationships Among Isoforms. Characterization of a Gene for  
 Starch-Branching Enzyme Iia from the Wheat D Genome Donor Aegilops  
 tauschii  
 Plant Physiol. 125 (3), 1314-1324 (2001)  
 MEDLINE 21140316  
 PUBMED 11244112  
 REFERENCE  
 2 (bases 1 to 11475)  
 Rahman, S., Regina, A., Li, Z., Sharon, A. and Morell, M. K.  
 Direct Submission  
 Submitted (18-JAN-2001) Plant Industry, CSIRO, Clunies Ross,  
 Canberra, ACT 2601, Australia  
 FEATURES  
 Source  
 location/Qualifiers  
 1. 11475  
 /organism="Aegilops tauschii"  
 /db\_xref="taxon:37682"  
 /chromosome="2"  
 join (<1204, .1336,1664, .1761,2038, .2279,2681, .2779,  
 2949, .2991,3144, .3203,3539, .3619,3703, .3819,4105, .4188,  
 4818, .4939,5115, .5234,6209, .6338,6427, .6537,6739, .6867,  
 7447, .7550,8392, .8536,9556, .9703,9839, .9939,10116, .10193,  
 10395, .10500,10928, .11002,11092, .11475)  
 /product="starch branching enzyme Iia"  
 join (1204, .1336,1664, .1761,2038, .2279,2681, .2779,  
 2949, .2991,3144, .3203,3539, .3619,3703, .3819,4105, .4188,  
 4818, .4939,5115, .5234,6209, .6338,6427, .6537,6739, .6867,  
 7447, .7550,8392, .8536,9556, .9703,9839, .9939,10116, .10193,  
 10395, .10500,10928, .11002,11092, .11175)  
 /note="SBEIIa"  
 /codon\_start=1  
 /product="starch branching enzyme Iia"  
 /protein\_id="AAK26821.1"  
 /db\_xref="GI:13447950"

BASE COUNT 3093 a 2253 c 2423 g 3689 t 11 others  
 ORIGIN





Query Match 13.6%; Score 136.4; DB 8; Length 2726;  
 Best Local Similarity 98.7%; Pred. No. 8.3e-20;  
 Matches 148; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 43 aggatgtatgatattcatgagctctggaatagg-cttcaactcttcgcatgacgtgagca 101  
 |||||||  
 DB 1787 AGGATGTATGATATTCAATGAGCTCTGGATAGAGCCTTCAACTCCTCCATGTATGCGGCA 1846  
 |||||||

QY 102 tagcatataataaagatcagagctgtcaccatgggttaggtggtgaagctactcta 161  
 |||||||  
 DB 1847 TAGCATTACATTAATAAGATCAGGCTTGTCACCATGGGTTAGGTGATGAAGGCTATCTTA 1906  
 |||||||

QY 162 actcatgggaaatgagttcgagcatcccg 191  
 |||||||  
 DB 1907 ACTTCATGGGAATGAGTTGGGCATCTG 1936  
 |||||||

RESULT 5  
 AF286319 2970 bp mRNA linear PLN 30-OCT-2000  
 LOCUS Triticum aestivum starch branching enzyme 2 (Sbe2) mRNA, complete  
 DEFINITION cds  
 ACCESSION AF286319  
 VERSION AF286319.1 GI:11037533  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 2970)  
 McCue,K.F., Hurman,W.J., Tanaka,C.K. and Anderson,O.D.  
 Starch Branching Enzymes Sbe1 and Sbe2 from Wheat (Triticum  
 aestivum cv. Cheyenne): Molecular Characterization, Developmental  
 Expression, and Homolog Assignment by Differential PCR  
 Unpublished  
 2 (bases 1 to 2970)  
 McCue,K.F. and Anderson,O.D.  
 Direct Submission  
 Submitted (11-JUL-2000) United States Department of Agriculture,  
 Agricultural Research Service, 800 Buchanan Street, Albany, CA  
 94710-1105, USA

FEATURES  
 source  
 1..2970  
 Location/Qualifiers  
 /organism="Triticum aestivum"  
 /cultivar="Cheyenne"  
 /db\_xref="taxon:4565"  
 /tissue\_type="endosperm"  
 1..2970  
 /gene="Sbe2"  
 149..2620  
 /EC\_number="2.4.1.81"  
 /note="glucosyltransferase; alpha-1,4-glucan;  
 alpha-1,4-glucan-6-glucosyltransferase"  
 /codon\_start=1  
 /product="starch branching enzyme 2"  
 /protein\_id="AAC27623.1"  
 /db\_xref="GI:11037534"  
 /translation="MATFAVSGATLGVARPAGAGGGLPRSGSERRGVDLPSSLARK  
 KDSRAVLSRAAPGKVLVPDGEDSLASPAQPELQIPEDI EGTAAVNMGTAAEK  
 LESSPPTQGIETITDGYTKGKELVGEKPRVYKPDGQKIYEIDPLKDFRSHLD  
 YRSYRIRIRAIIDHGGLEAFSGYKELGFTTSAEGITTYREMAPGAHSALVGDPN  
 NMNPAADMTDIDYGVWEITFLPNNADGSPALPHGSRVATRMOTPSGVADSIAMTKFS  
 VOAPGEIPFNGLIYDPPPEEKYVFOHPQPKPRESLRITESHIGMSSPEPKINSTANFR  
 DEVLPRIKRLGNAVOIMAIQESHYSYASFGYHVNFPAPSSRFGTPEDLSLIDRAHE  
 LGLVLAMDIVSHSSNNITLDGLNGFDGTDTHYFHGPRGHMMMDSRLEFNYSWVLR  
 FLASNAWMLLEKYGDFRFGVTSMTYHHLQMTFGNNGEYGFADVDAAVYLM  
 LVNDLIGHLPDVAISGDVSGMPTFCIPVPGGVGPDYRLHMAVAADKMIELKQSDS  
 SKMGDIYHTLTPNRMLEKCYTAAASHDOALVGDCTIAFWMLDKMYDPMALDRSTP  
 RIDRGLAIHKMIRLVTYKGLGGEGYLNFGNENRGNHEWIDFPRGPOTLFGVLPNNN  
 SYDKCRKRPDGLADFLRYHGMQERDQAMQHLLEKYGFTSHQYVSKRHEEDKYLIF  
 ERGDLVFEVFNHMSNPFEDYRVGCSRPQKYVALDSDALFEGFSRLDHDVYFTTEH

Query Match 13.6%; Score 136.4; DB 8; Length 2970;  
 Best Local Similarity 98.7%; Pred. No. 8.2e-20;  
 Matches 148; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 43 aggatgtatgatattcatgagctctggaatagg-cttcaactcttcgcatgacgtgagca 101  
 |||||||  
 DB 1977 AGGATGTATGATATTCAATGAGCTCTGGATAGAGCCTTCAACTCCTCCATGTATGCGGCA 2036  
 |||||||

QY 102 tagcatataataaagatcagagctgtcaccatgggttaggtggtgaagctactcta 161  
 |||||||  
 DB 2037 TAGCATTACATTAATAAGATCAGGCTTGTCACCATGGGTTAGGTGATGAAGGCTATCTTA 2096  
 |||||||

QY 162 actcatgggaaatgagttcgagcatcccg 191  
 |||||||  
 DB 2097 ACTTCATGGGAATGAGTTGGGCATCTG 2126  
 |||||||

RESULT 6  
 TASBA2 2970 bp mRNA linear PLN 27-FEB-1997  
 LOCUS T. aestivum mRNA for starch branching enzyme II.  
 DEFINITION Y11282  
 ACCESSION Y11282.1 GI:1885343  
 VERSION Y11282.1 GI:1885343  
 KEYWORDS  
 1,4-alpha-glucan branching enzyme; sbe2 gene; starch branching  
 enzyme II.  
 SOURCE  
 Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 2970)  
 Nair,R.B., Baga,M., Scoles,G.J., Kartha,K.K. and Chibbar,R.N.  
 Isolation, characterization and expression analysis of a starch  
 branching enzyme II cDNA from wheat  
 Plant Sci. In press  
 2 (bases 1 to 2970)  
 Chibbar,R.N.  
 Direct Submission  
 Submitted (14-FEB-1997) R.N. Chibbar, Plant Biotechnology  
 Institute, National Research Council of Canada, 110 Gymnasium  
 Place, Saskatoon, Saskatchewan S7N 0W9, CANADA

FEATURES  
 source  
 1..2970  
 Location/Qualifiers  
 /organism="Triticum aestivum"  
 /cultivar="Fielder"  
 /db\_xref="taxon:4565"  
 /clone="PRN33"  
 /dev\_stage="kernels 12 days post anthesis"  
 151..2948  
 /gene="Sbe2"  
 151..312  
 /gene="Sbe2"  
 151..2622  
 /EC\_number="2.4.1.18"  
 /codon\_start=1  
 /product="1,4-alpha-glucan branching enzyme II"  
 /protein\_id="CA472154.1"  
 /db\_xref="GI:1885344"  
 /db\_xref="SPTREMBL:P93691"  
 /translation="MATFAVSGATLGVARPAGAGGGLPRSGSERRGVDLPSSLARK  
 KDSRAVLSRAAPGKVLVPDGEDSLASPAQPELQIPEDI EGTAAVNMGTAAEK  
 LESSPPTQGIETITDGYTKGKELVGEKPRVYKPDGQKIYEIDPLKDFRSHLD  
 YRSYRIRIRAIIDHGGLEAFSGYKELGFTTSAEGITTYREMAPGAHSALVGDPN  
 NMNPAADMTDIDYGVWEITFLPNNADGSPALPHGSRVATRMOTPSGVADSIAMTKFS  
 VOAPGEIPFNGLIYDPPPEEKYVFOHPQPKPRESLRITESHIGMSSPEPKINSTANFR  
 DEVLPRIKRLGNAVOIMAIQESHYSYASFGYHVNFPAPSSRFGTPEDLSLIDRAHE  
 LGLVLAMDIVSHSSNNITLDGLNGFDGTDTHYFHGPRGHMMMDSRLEFNYSWVLR

FLSNARWMLSEYFDFGRPDGVTSMATYTHHGLQMTTGATNGEYFGPATDVAVYILM  
LVNDLHGLHPDASIGEDVSGMPTFECIPVDDGVGLDYRLHMAVADKWLLELKQSE  
SMKGDIVYHTLITNRMLSEKCYTAESHQALVGDKTIAFMLMDMDYFADLDRPSTP  
RIDRIGIALHKMIRLVMTGLGEGYLANFMNEFGHEMIDEPGDPOTLPTGVLGNNN  
SYDCRRRFDLGDADFLRYHGMQFDOAMOLEEKYGMTESEHOYVSRKHEDVYLIE  
ERGLDVFVFNFMHNSNPFEDYRVGCSRGKTKVALDSDALEGGSRIDHDVDTTTEH  
PHDNRPSFSVYTTSTRVAVYALTE"  
313..2619  
/gene="sbe2"  
/EC\_number="2.4.1.18"  
/product="1,4-alpha-glucan branching enzyme II"  
polyA\_signal  
2879..2884  
/gene="sbe2"  
2930..2935  
/gene="sbe2"  
polyA\_site  
2948  
/gene="sbe2"  
BASE COUNT 782 a 640 c 775 g 773 t  
ORIGIN

Query Match 13.6%; Score 136.4; DB 8; Length 2970;  
Best Local Similarity 98.7%; Pred. No. 8.2e-20;  
Matches 148; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 43 aggaatgatagtatcctgctgagatagg-cttcaactcttcgcatgacgtggca 101  
|||||  
DB 1979 AGGATATGATGATTTTCATGCGCTCGATAGGCGCTTCAACTCGCATGTATGCTGGCA 2038  
|||||  
QY 102 tagcattacataaatagtacagcgtctgcacacatgggttagtggtgaagcatctcta 161  
|||||  
DB 2039 TAGCATTTACATTAATAATGATACAGGCTTGTCACCATGCGTTTAGTGGTGAAGCTATCTTA 2098  
|||||

QY 162 actcatatggaatgagttggcgcctcctg 191  
|||||  
DB 2099 ACTTCATGGGAATGACGTTTGCGCATCTCTG 2128  
|||||

RESULT 7  
TAU66376 2853 bp mRNA linear PLN 18-OCT-1996  
LOCUS  
DEFINITION 6-alpha-D-(1,4-alpha-D-glucanotransferase mRNA, complete cds.  
ACCESSION U66376  
VERSION U66376.1 GI:1620661  
KEYWORDS  
SOURCE  
ORGANISM  
Triticum aestivum  
wheat.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Pooideae; Triticeae; Triticum.  
1 (bases 1 to 2853)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (08-AUG-1996) University of Hamburg, Institute of General  
Botany, Centre of Applied Molecular Biology, AMP II, Ohnhorststr.  
18, Hamburg 22609, Germany  
Location/Qualifiers  
1..2853  
/organism="Triticum aestivum"  
/cultivar="T.A. Florida"  
/db\_xref="taxon:4565"  
/tissue\_type="kernels 21 DAP"  
313..2502  
/EC\_number="2.4.1.18"  
/note="branching enzyme"  
/codon\_start=1  
/product="1,4-alpha-D-glucan  
6-alpha-D-(1,4-alpha-D-glucanotransferase"  
/protein\_id="AA017086.1"  
/db\_xref="GI:1620662"  
/translation="MTGGTAEXKLOSSPPGQIVETLTIDGYTKGVKKEIVGKPRVYVK  
PGGQKRIYEIDPLTKDRSHLDYRKYREKRIIRAIDQHEGGLAEFSGYKELGTFTRSA

EGTIRREMAFGAHSALVGVGNMKNPADAMTRDDYCYWELFLPNNADGSSALPHGSR  
VKIRMDTPSSVSKOSISAMIKFVQAPPEIPFNIGTYDPEEKYVQHPQKRPESLR  
IYESHIISSPEPEKINSYANFROVEYLRIRLGINAVQIAIOESHYSAGCYHYNF  
FASSRRGTEDLKLILDRALHEGLILVLMIDIVSHSSNNTLDGJLNGDGTDTYFHGG  
PRGHMMWDRSLFNYGSMVEYLRFLSNARWMLSEYFDFGRPDGVTSMATYTHHGLQMT  
FTGNYGCEPFPATDVAVYILMVLVNDLHGLYPAVSGIGDVSGMPTFECIPVDDGVG  
PDTRLHMAVADKWLLELKQSEBKRMDIYHTLTNRKLEKCYTAESHQALVGDKT  
IAFMLMDMDYFADLDRPSTPRIDRIGIALHKMIRLVMTGLGEGSYLANFMNEFGHE  
WIDFPRGPOTLPTGKVLPGNNNNYDKCRRRFDLGDADFLRYHGMQFDOAMOLEEKY  
GFMTSEHOYVSRKHEDKVIIFERGDVLFVFNFMHNSNPFEDYRVGCSRGKTKVALDSD  
DDALFGEFSRLHDVDVETTEHPDNRPSFSVYTTSTRVAVYALTE"  
BASE COUNT 776 a 585 c 746 g 746 t  
ORIGIN

Query Match 13.5%; Score 134.8; DB 8; Length 2853;  
Best Local Similarity 98.0%; Pred. No. 1.8e-19;  
Matches 147; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 43 aggaatgatagtatcctgctgagatagg-cttcaactcttcgcatgacgtggca 101  
|||||  
DB 1859 AGGATATGATGATTTTCATGCGCTCGATAGGCGCTTCAACTCGCATGTATGCTGGCA 1918  
|||||  
QY 102 tagcattacataaatagtacagcgtctgcacacatgggttagtggtgaagcatctcta 161  
|||||  
DB 1919 TAGCATTTACATTAATAATGATACAGGCTTGTCACCATGCGTTTAGTGGTGAAGCTATCTTA 1978  
|||||  
QY 162 actcatatggaatgagttggcgcctcctg 191  
|||||  
DB 1979 ACTTCATGGGAATGAGTTTGCGCATCTCTG 2008  
|||||

RESULT 8  
AF064560 2554 bp mRNA linear PLN 03-NOV-1998  
LOCUS  
DEFINITION Hordeum vulgare Bomi Starch branching enzyme IIA (sbeIIa)  
ACCESSION AF064560  
VERSION AF064560.1 GI:3822019  
KEYWORDS  
SOURCE  
ORGANISM  
barley.  
Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Pooideae; Triticeae; Hordeum.  
1 (bases 1 to 2554)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
98404232  
2 (bases 1 to 2554)  
Sun C., Sathish P., Ahlandsberg S. and Jansson C.  
The two genes encoding starch-branching enzymes IIA and IIB are  
differentially expressed in barley  
Plant Physiol. 118 (1), 37-49 (1998)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (11-MAY-1998) Stockholm University, Biochemistry,  
Stockholm S-10691, Sweden  
Location/Qualifiers  
1..2554  
/organism="Hordeum vulgare"  
/cultivar="Bomi"  
/db\_xref="taxon:4513"  
/chromosome="2"  
1..2554  
/gene="sbeIIa"  
7..2211  
/gene="sbeIIa"  
/EC\_number="2.4.1.18"  
/codon\_start=1  
/product="starch branching enzyme IIA"  
/protein\_id="AAC69753.1"  
/db\_xref="GI:3822020"  
/translation="MAEVNMTGGAAGKLESSPPTGGIAETITDGYTKGVKKEIVGKPRVYVK  
QVVPKRDGQKRIYEIDPLTKDRSHLDYRKYREKRIIRAIDQHEGGLAEFSGYKELGTFTRSA

FTSAKITYREMAPGAHSALVGDPMNPNMADPMTRDXYGWEIFELPNADGSPAI  
PHGRVKTRMDTPSGVKDSISAMTKFSQACETIPNCTIYDPRPEEYVYQHOPRK  
PESLRITESHGMSPEPKINSYIRFDEVLPRIKGLYNVQIMAOEHSYVASFQ  
HVTNFEAPSSREFPEDLSLIDRAHEGLVLMIVHSHSNNTLDOINFDSTDT  
YFHGPRHMMMDPSRLFNKSGSEVRLISNARWLEXYKFDFRFGVYSMTYTH  
GIOMTFNGYGEYGFATVDVAVYLMVNDLIGLVPDAYSIGEDVSGMPTFCIPV  
DGGVDFRLHMAVADKWIILLKQSDSKMGDIIVHTLRNMLEKCTVYAESDQAL  
VGDKTIAFWLMDKMDYMDALDRSPTRIDGIALHKIRLVTMGIGEGEGLNMGNE  
FGHEMIDFPRGPTLPTGKYLPGNNNSYDKCRFRFDGDDFLRYRMQFEDAMOH  
LEERYGWTSEHQYVSRKHEDEKVIIFERGLVVEFNFHWSNSKRDYRVCSGKGYK  
VALSDALFEFGFSRLDHDVYFTTEHPHDNRPSFSYVTPSRVAYVALNE"

BASE COUNT 731 a 494 c 632 g 697 t

ORIGIN

Query Match 13.0%; Score 130; DB 8; Length 2554;  
Best Local Similarity 96.0%; Pred. No. 2.1e-18;  
Matches 144; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 43 aggaatgataattcattcgtcgtcgtgtag-gctcaactcttcgattcgtgca 101  
|||||  
Db 1568 AGGATATGATGATTTTCATGCTCTGGATAGACCTTCAACCCCTCGCATGTATCGGGCA 1627  
|||||

Qy 102 tagcatatcataaaatgataagctgtcaccatggttagtggtgaagctatccta 161  
|||||  
Db 1628 TAGCATTTACATTAATAATGATCAGCTTGTCACCATGGGTTAGTGGCGAAGCTATCTTA 1687  
|||||

Qy 162 actcaatgggaatgattggcgcctctg 191  
|||||  
Db 1688 ATTTCATGGGAATGAGTTGGGCATCTCG 1717  
|||||

RESULT 9

LOCUS 2M065948 2795 bp mRNA linear PLN 21-AUG-1997  
DEFINITION Zea mays starch branching enzyme Iia (Sbe2a) mRNA, partial cds.  
ACCESSION U65948  
VERSION U65948.1 GI:2340107  
KEYWORDS  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 2795)  
AUTHORS Geo,M., Fisher,D.K., Kim,K.N., Shannon,J.C. and Gulltman,M.J.  
TITLE Independent genetic control of maize starch-branching enzymes Iia  
and Iib. Isolation and characterization of a Sbe2a cDNA  
JOURNAL Plant Physiol. 114 (1), 69-78 (1997)  
MEDLINE 97303618  
AUTHORS 2 (bases 1 to 2795)  
JOURNAL Geo,M., Fisher,D.K., Kim,K.-N., Shannon,J.C. and Gulltman,M.J.  
TITLE Direct Submission  
REFERENCE Submitted (02-AUG-1996) Hort.,Biotech.Inst., Penn State University,  
AUTHORS 519 Wartik, University Park, PA 16802, USA  
JOURNAL  
FEATURES  
source Location/Qualifiers  
1. 2795  
/organism="Zea mays"  
/strain="B73"  
/db\_xref="taxon:4577"  
/note="inbred"  
<1..2795  
/gene="Sbe2a"  
<1..2446  
/gene="Sbe2a"  
CDS  
/EC\_number="2.4.1.18"  
/function="formation of alpha-1-6 glucosidic linkage in  
starch biosynthesis"  
/note="starch branching enzyme Iisozyme SBEIIa"  
/codon\_start=2  
/product="starch branching enzyme Iia"  
/protein\_id="AAB67316.1"  
/db\_xref="GI:2340108"

/translation="DLPSEVLRFRKDAFSRVLSCAGAPGKVLVPGGSDLLSSABPV  
VDTPPEBEQIPEALVLYEKSSSPQTSSANAEESSVEAEERPELSEVIGVGTGCT  
KIDGAGIKAKPLVEERPRVLPDGDQRIEIDPMLEGFEGHLDYRSEYKRLAAI  
DQHGAGDAEPRGSEYKLGFTNSAGVITREMAAPAYSAALVGDPMNPNADAAARNE  
YGVMEIFELPNADGSPAI PHGRVK IIMDTPSGKDSIPAIKFSVOAPGRIPYNGIY  
YDPEEKEVYVYKHOPRKSLRIYESHVGMSPEPKINTYANRDEVLPRIKGLYN  
AVOIMAOEHSYVAFVGHVNTNFEAPSSREFPEDLSLIDRAHEGLVLMIVHSH  
SNNTLDOINFDSTDTTHYFPGGRHMMMDPSRLFNKSGSEVRLISNARWLEXY  
KFDFRFGDGTSMKTHHGLVYITGNTGETFGATVDVAVYLMVNDLIRGLTPEA  
VSIQEDVSGMPTFCIPVODGGVGFYDRIHMAVDPKWIILLKQSDSEYEMGDIVHTLN  
RRLMEKCVYCESHDOALVGDKTIATFIMLMDKMDYMDALDRSPTRIDGIALHKMR  
LVTMGIGEGEGLNMGNEFGHEMIDFPRGPTLPTGKYLPGNNNSYDKCRFRFDG  
ADLYRYGMQFEDAMOHLEEGYEFEMTSDEHSYVSRKHEDEKVIIFERGLVVEFNFH  
SNSYEDRYVCGCFKPKYKIVDSDDGLFGGFSRLDHDVYFTTEHPHDNRPSFSYV  
PSRTAVYIAPGAEDE"

BASE COUNT 766 a 563 c 709 g 757 t

ORIGIN

Query Match 12.2%; Score 122; DB 8; Length 2795;  
Best Local Similarity 92.7%; Pred. No. 1.1e-16;  
Matches 139; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

Qy 43 aggaatgataattcattcgtcgtcgtgtag-gctcaactcttcgattcgtgca 101  
|||||  
Db 1791 AGGATATGATGATTTTCATGCTCTGGACAGGCTTCAACGCCCTCGCATGTATCGGGCA 1850  
|||||

Qy 102 tagcatatcataaaatgataagctgtcaccatggttagtggtgaagctatccta 161  
|||||  
Db 1851 TAGCATTTACATTAATAATGATGAGTTAGGAGGTGAGAGGCTATCTTA 1910  
|||||

Qy 162 actcaatgggaatgattggcgcctctg 191  
|||||  
Db 1911 ATTTCATGGGAATGAGTTGGGCATCTCG 1940  
|||||

RESULT 10

LOCUS E14724 2364 bp DNA linear PAT 28-JUL-1999  
DEFINITION Rice mRNA for branching enzyme-4 mature peptide.  
ACCESSION E14724  
VERSION E14724.1 GI:5709407  
KEYWORDS JP 1998004970-A/2.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartioideae; Oryzaeae; Oryza.  
REFERENCE 1 (bases 1 to 2364)  
AUTHORS Baba,T., Kawasaki,T. and Ichikawa,N.  
TITLE NEW RICE GENE FOR STARCH-BRANCHING ENZYME  
JOURNAL Patent: JP 1998004970-A 2 13-JAN-1998;  
COMMENT MITSUI GYOSAI SHOKUDUTSU BIO KENYUSHO:KK, MITSUI PETROCHEM IND LTD  
Oryza sativa (rice)  
PN JP 1998004970-A/2  
PD 13-JAN-1998  
PF 24-JUN-1996 JP 1996162983  
PI BABA TADASHI, KAWASAKI TSUTOMU, ICHIKAWA NORIO PC  
C12N15/09,A01H5/00,C07H21/04,C12N5/10,C12N9/10,C12N15/09, PC  
C13R1:91),  
PC (C12N5/10,C12R1:91),(C12N9/10,C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FH source 1..2364  
FT /organism="Oryza sativa"  
FT /cultivar="Nihonbare"  
FT /clone\_lib="Rise Immature Seed Lambda gtII FT  
FT /tissue\_type="Immature Seed"  
FT /clone="pRBA1"  
FT mat\_peptide 1..2364

FEATURES FT /product='branching enzyme-4'.  
source 1. .2364  
/db\_xref='taxon:4530'  
BASE COUNT 670 a 454 c 600 g 640 t  
ORIGIN

Query Match 11.9%; Score 118.8; DB 6; Length 2364;  
Best Local Similarity 91.3%; Pred. No. 5.7e-16;  
Matches 137; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 43 agagatgatgatattcattgctctgtagatg-gcttaactcttcgattgtagcga 101  
|||||  
Db 1721 AGGATATGATGATTTATGAGCTGTACACAGACCTTCAACCGCGCATGATGCTGGGA 1780

QY 102 tagcattacataaataatgataagctgtcaccatgggttagtgatgtagaagcattccta 161  
|||||  
Db 1781 TAGCATTACATTAATAATGATTAAGCTGTACACAGAGCGGCGCATGATCTTA 1840

QY 162 acttcattggaatgagttcgggcatcctg 191  
|||||  
Db 1841 ATTTCATGGGAATGAGTTGGGCATCCTG 1870

RESULT 11  
LOCUS E14723 3015 bp DNA linear PAT 28-JUL-1999  
DEFINITION Rice mRNA for branching enzyme-4, complete cds.  
ACCESSION E14723.1 GI:5709406  
VERSION JP 1998004970-A/1.  
KEYWORDS Oryza sativa.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 3015)  
Baba,T., Kawasaki,T. and Ichikawa,N.  
NEW RICE GENE FOR STARCH-BRANCHING ENZYME  
Patent: JP 1998004970-A 1 13-JAN-1998;  
MITSUI GIOSAI SHOKUBUTSU BIO KENKUSHO:KK, MITSUI PETROCHEM IND LTD  
COMMENT OS Oryza sativa (rice)  
PN JP 1998004970-A/1  
PD 13-JAN-1998  
PF 24-JUN-1996 JP 1996162983  
PI BABA TADASHI, KAWASAKI TSUTOMU, ICHIKAWA NORIO PC  
C12N15/09,A01H5/00,C07H21/04,C12N5/10,C12N9/10,C12N15/09, PC  
C12R1:911,  
PC (C12N5/10,C12R1:91),(C12N9/10,C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
FH Key  
FH Location/Qualifiers  
FT source 1. .3015  
FT /organism='Oryza sativa'  
FT /cultivar='Nihondare'  
FT /clone\_lib='Rice Immature Seed Lambda gIII FT  
FT CDNA library'  
FT /tissue\_type='Immature seed'  
FT /clone='pRB41'  
FT CDS 129. .2654  
/product='branching enzyme-4' FT sig\_peptide

FEATURES  
source 1. .3015  
/organism='Oryza sativa'  
/db\_xref='taxon:4530'  
BASE COUNT 796 a 606 c 819 g 794 t  
ORIGIN

Query Match 11.9%; Score 118.8; DB 6; Length 3015;  
Best Local Similarity 91.3%; Pred. No. 5.5e-16;  
Matches 137; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 43 agagatgatgatattcattgctctgtagatg-gcttaactcttcgattgtagcga 101  
|||||  
Db 2008 AGGATATGATGATTTATGAGCTGTACACAGACCTTCAACCGCGCATGATGCTGGGA 2067

QY 102 tagcattacataaataatgataagctgtcaccatgggttagtgatgtagaagcattccta 161  
|||||  
Db 2068 TAGCATTACATTAATAATGATTAAGCTGTACACAGAGCGGCGCATGATCTTA 2127

QY 162 acttcattggaatgagttcgggcatcctg 191  
|||||  
Db 2128 ATTTCATGGGAATGAGTTGGGCATCCTG 2157

RESULT 12  
LOCUS AB023498  
DEFINITION Oryza sativa mRNA for starch branching enzyme rbe4, complete cds.  
ACCESSION AB023498  
VERSION AB023498.1 GI:5689137  
KEYWORDS starch branching enzyme rbe4.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (sites)  
Mizuno,K., Tachibana,M., Kobayashi,E., Kawasaki,T., Funane,K.,  
Kobayashi,M. and Baba,T.  
Molecular cloning and expression analysis of a novel member of  
starch branching enzyme isoform in developing rice seeds  
Unpublished (1999)  
2 (bases 1 to 3015)  
Mizuno,K. and Baba,T.  
Direct Submision  
Submitted (09-FEB-1999) Kouichi Mizuno, University of Tsukuba,  
Institute of Agricultural and Forest Engineering; 1-1-1 Tennoudai,  
Tsukuba, Ibaraki 305-8572, Japan  
(E-mail:koumomo@tsukuba.ac.jp, Tel:81-298-53-4656,  
Fax:81-298-53-2203)  
FEATURES  
source 1. .3015  
/organism='Oryza sativa'  
/db\_xref='taxon:4530'  
/tissue\_lib='developing seeds'  
129. .2654  
/gene='RBE4'  
/gene='RBE4'  
129. .2654  
/gene='RBE4'  
/codon\_start=1  
/product='starch branching enzyme rbe4'  
/protein\_id='BA82828.1'  
/db\_xref='GI:5689138'  
/translation='MASFAVSGARLGVNAGGGGGGGGGGPPANSGVDLPVLFRRKD  
SFSRGVNSGAGAGKVLVPGGSDLLSAPVEVPEOPEEQQIDDDKVKFEDEE  
EIPAVAEASIKVAEDKLESEVIODIEEVNTEVDEPVEEKPRVIEPBGGO  
KIYOIDPMLIEGFNNHLDVRYSEVKRRAAIDOHGGGLDAFSEKELGFRSAGITY  
REMAFGAQSALVGDPENNNNPNADVTWRENYGSESLPNMAGSGAIPRGSGVKT  
DTPSGVKDSTPAWIKFAVDAFCETPYNGITYDPEEKEVYFQHPKRRNSLRTYSH  
IGMSSEPKINTYANFRDVLPRIKLIGYNAVOIMAIQESHYYASFGYHVTNEFAPS  
RFGTPEDLKSLIDKAHELGLVLMADIHSHASNNITDINGNGEGDTVHYHGGPRGH  
WMMDSRLLFNYSWSEVYRLLSNARWMLLEEKFGDFPDGVTSMYTHHGLQVAFGTNY  
GEYFGATGIVDAVYVIMLVNDLIHGLYPRAVAIAGEVSGAPFCIPVODGGGFTFRL  
HNAVPKATIEFLKQSDPEYKMGDIYVITLNNRRSEKCYAAESHDAIAGDGTIAFML  
MDKMDTDEMALDREPSTPRIDRGIALHKMLRLVTWGLGSGYLNNFGNENFGHPWIDFP  
RGOYSLPNGSVLPNNYSPDKCRRPDLADADLRYHAGQEEFOAQMLHEEYXGWTSS  
EHQYIISRKHEDEKVIIFERGDIVFVNFPMWSNYSFPRYRGCLKPGFYKKIIVLSDGLF  
GGPSRLDHAEEYFTADWPHDNRCPSPSVYTPSRTAVVYALTED"

BASE COUNT 796 a 606 c 819 g 794 t  
ORIGIN



Query Match 11.9%; Score 118.8; DB 8; Length 3015;  
 Best Local Similarity 91.3%; Pred. No. 5.5e-16;  
 Matches 137; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 43 aggaatgatgatgattcgaagctctgtagatag-gcttcaactcttcgcatgtagcgca 101  
 |||||||  
 DB 2008 AGGATATGATGATTTTATAGGCTCTAGACAGACCTTCAACACTCCGATGATGCGGA 2067  
 |||||||

QY 102 tagcatatacaaaatgcatcaggtctgtcaccatgggttaggtgtagaagctatctta 161  
 |||||||  
 DB 2068 TAGCATTACATTAATAAGATTAGGCTGTGCACCATGGGCTTAGAGGCGAAGGCTATCTTA 2127  
 |||||||

QY 162 actcatgggaaatgagttgggcatcccg 191  
 |||||||  
 DB 2128 ATTTCATGGGAATGAGTTGGCATCCTG 2157  
 |||||||

RESULT 13  
 ARI06495 2725 bp DNA linear PAT 14-FEB-2001  
 LOCUS ARI06495 Sequence 14 from patent US 6107060.  
 DEFINITION ARI06495  
 ACCESSION ARI06495  
 VERSION ARI06495.1 GI:12821025  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2725)  
 AUTHORS Keeling, P. and Guan, H.  
 TITLE Starch encapsulation  
 JOURNAL Patent: US 6107060-A 14 22-AUG-2000;  
 FEATURES Location/Qualifiers  
 source 1..2725  
 BASE COUNT 727 a 534 c 715 g 749 t  
 ORIGIN

Query Match 11.1%; Score 110.8; DB 6; Length 2725;  
 Best Local Similarity 88.0%; Pred. No. 3.1e-14;  
 Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 43 aggaatgatgatgattcgaagctctgtagatag-gcttcaactcttcgcatgtagcgca 101  
 |||||||  
 DB 1847 AGGATATGATGATTTTATAGGCTCTAGACAGACCTTCAACACTCCGATGATGCGGA 1906  
 |||||||

QY 102 tagcatatacaaaatgcatcaggtctgtcaccatgggttaggtgtagaagctatctta 161  
 |||||||  
 DB 1907 TAGCATTACATTAATAAGATTAGGCTGTGCACCATGGGCTTAGAGGCGAAGGCTATCTTA 1966  
 |||||||

QY 162 actcatgggaaatgagttgggcatcccg 191  
 |||||||  
 DB 1967 ATTTCATGGGAATGAGTTGGCATCCTG 1996  
 |||||||

RESULT 14  
 MZEGUOCTRN 2725 bp mRNA linear PLN 01-JUN-1994  
 LOCUS MZEGUOCTRN  
 DEFINITION Corn starch branching enzyme II mRNA, complete cds.  
 ACCESSION L08065  
 VERSION L08065.1 GI:168482  
 KEYWORDS 1,4-alpha-glucan branching enzyme; amylo-transglucosylase;  
 glucanotransferase; starch branching enzyme II.  
 SOURCE Zea mays cDNA to mRNA.  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 2725)  
 REFERENCE Fisher, D.K., Boyer, C.D. and Hannah, L.C.  
 AUTHORS  
 JOURNAL  
 TITLE  
 REFERENCE  
 AUTHORS  
 TITLE

JOURNAL Plant Physiol. 102, 1045-1046 (1993)  
 MEDLINE 94105320  
 FEATURES Location/Qualifiers  
 source 1..2725  
 /organism="Zea mays"  
 /cultivar="W64Ax182E"  
 /db\_xref="taxon:4577"  
 /tissue\_type="endosperm"  
 /dev\_stage="29 days post pollination"  
 91..264  
 CDS 91..2490  
 /EC\_number="2.4.1.18"  
 /codon\_start=1  
 /product="starch branching enzyme II"  
 /protein\_id="AAA18571.1"  
 /db\_xref="GI:168483"  
 /translation="MAFRVSGAVLGAARAPRLTGGEGSLVFRHTGLFLRGARVGC  
 SGTGAMRAAARAAVAVPEGEENDGLASRADSAQFOSDELEVPDISERTTCAGVAD  
 AGLNRVRYVPPSDGOKIFQIDPMLOGYKYLERYSLYRIRSDIDEBEGLEAFS  
 RYERKGFMAAEGLITREWARFASALVGVNMMMDPADMSKNEGVWEIFLPMN  
 ADGTSPIPHGSRVVKVRMDIPSGIKDSIPAMIKYSQABEIPDCTIYDDPEVYVYF  
 RHAQPKRPSLRIYETHVGMSSPEPRINTYVNRDEVLPRIKLSTNAQIMATDEHS  
 YYSGEGYHTVNFAPSRFGTEDELSLRDHEGLVLMVSHASSNTLDGLNG  
 FGDGTHYFHSQPRGHMMMDRLFNYGMEVLRFLLSNARMMLREYKFDGFRDGV  
 SMAYTHHGLQVTFGTNENYFGEATDVDAVVYVLMVNDLHGLYPEAVTIGEDVGMP  
 PFLPVHOGVGEDVRYMMAVADKMTDILKOSDETMMKGDIVHTLJNRMLEKCYTVA  
 ESHDVALVGDKTIAFWLMDKMTDFEMALDREPTPTIDGIALHKMIRLITMGLGEGY  
 LNFNGENRFGHPWIDFPRGPORLPFGSKFTPGNNNSYDKCRFRFDGADAVLYRHMOE  
 FDOAMQHLQKYEFTSDHQYISRHHEEDKLVFERGDLVFVFNHCNNSYFDYRIGC  
 RKPQVYKVLVDSADGLFGGFSRIHHAHEHFTADCSHDNRPYSFSVYTSRTCVVYAVP  
 E"

mat\_peptide 265..2487  
 /product="starch branching enzyme II"  
 BASE COUNT 727 a 534 c 715 g 749 t  
 ORIGIN

Query Match 11.1%; Score 110.8; DB 8; Length 2725;  
 Best Local Similarity 88.0%; Pred. No. 3.1e-14;  
 Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 43 aggaatgatgatgattcgaagctctgtagatag-gcttcaactcttcgcatgtagcgca 101  
 |||||||  
 DB 1847 AGGATATGATGATTTTATAGGCTCTAGACAGACCTTCAACACTCCGATGATGCGGA 1906  
 |||||||

QY 102 tagcatatacaaaatgcatcaggtctgtcaccatgggttaggtgtagaagctatctta 161  
 |||||||  
 DB 1907 TAGCATTACATTAATAAGATTAGGCTGTGCACCATGGGCTTAGAGGCGAAGGCTATCTTA 1966  
 |||||||

QY 162 actcatgggaaatgagttgggcatcccg 191  
 |||||||  
 DB 1967 ATTTCATGGGAATGAGTTGGCATCCTG 1996  
 |||||||

RESULT 15  
 AF451191 419 bp mRNA linear PLN 18-DEC-2001  
 LOCUS AF451191  
 DEFINITION Oryza sativa granule-bound starch synthase mRNA, partial cds.  
 ACCESSION AF451191  
 VERSION AF451191.1 GI:17902268  
 KEYWORDS  
 SOURCE Oryza sativa.  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoideae; Oryza.  
 1 (bases 1 to 419)  
 REFERENCE Lee, D.-S. and Hur, Y.  
 AUTHORS Expression and regulation of genes involved in carbohydrate  
 TITLE metabolism in rice  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 419)  
 AUTHORS Lee, D.-S. and Hur, Y.

TITLE Direct Submission  
JOURNAL Submitted (23-NOV-2001) Biology, Chungnam National University, 220  
Kungdong, Yuseong-Ku, Taejeon 365-764, Korea

FEATURES  
Source location/Qualifiers  
1. .419

/organism="Oryza sativa"  
/cultivar="Stejaree45"  
/db\_xref="taxon:4530"  
<1. .>419

/codon\_start=1  
/product="granule-bound starch synthase"

/protein\_id="AAL47851.1"

/db\_xref="GI:17902269"

/translation="LYPEAVAIGEDVSGMPTFCIPVODGVGFYRLHMAVPDKWIGL  
LKODEYWKMGDIYHTLTNRMSKCVYAESHOALYGDXTIAFWLMDKMDYDMAL  
DRPSTPRIDRIALHKMIRLVTMGIGGGYLNFMGNE"

BASE COUNT 115 a 71 c 110 g 123 t

ORIGIN

Query Match 10.8%; Score 107.8; DB 8; Length 419;

Best Local Similarity 90.6%; Pred. No. 1.8e-13;

Matches 126; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 43 aggatatgatgattcattcgtgctcgatag-gcttcaactcttcgattgacgtggca 101  
|||||  
DB 281 AGGATATGATGATTATTATGGCTTAGACAGACCTTCAACACCTCGCATGATCGTGGA 340  
QY 102 tagcattacataaagtcaggctgtcacatggttagtggtgaagctactta 161  
|||||  
DB 341 TAGCATTTACATTAATAATGATTAGCTGTGCACCATGGCTTAGAGGCGAAGGCTATCTTA 400  
QY 162 acttcattggaatgagtt 180  
|  
DB 401 ATTTCATGGGAATGAGTT 419

Search completed: September 12, 2002, 07:30:44  
Job time: 20832 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 08:21:51 : Search time 406.92 Seconds

(without alignments)  
4223.512 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_9500\_10500

Perfect score: 1001  
Sequence: 1 tccgacatgacacagcatga.....ctggagcaatagctttttg 1001

Scoring table: IDENTITY\_NUC  
Gapop.10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_032802:\*

- 1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 1001  | 100.0       | 11473  | 20    | AA34650     |
| 2          | 978   | 97.7        | 11475  | 22    | AAH78338    |
| 3          | 223.2 | 22.3        | 23449  | 21    | AA253393    |
| 4          | 136.4 | 13.6        | 2726   | 22    | AAH78337    |
| 5          | 133.4 | 13.3        | 979    | 21    | AA299942    |
| 6          | 130.2 | 13.0        | 977    | 21    | AA299947    |
| 7          | 130.2 | 13.0        | 984    | 21    | AA299946    |
| 8          | 118.8 | 11.9        | 3015   | 19    | AAV05639    |
| 9          | 114.2 | 11.4        | 1085   | 21    | AA299962    |

|    |       |      |      |    |          |
|----|-------|------|------|----|----------|
| 10 | 110.8 | 11.1 | 2087 | 18 | AA69737  |
| 11 | 110.8 | 11.1 | 2165 | 18 | AA69736  |
| 12 | 110.8 | 11.1 | 2640 | 19 | AAV70961 |
| 13 | 110.8 | 11.1 | 2665 | 18 | AA69729  |
| 14 | 110.8 | 11.1 | 2725 | 18 | AAV28737 |
| 15 | 106   | 10.6 | 2919 | 15 | AAQ73750 |
| 16 | 102.8 | 10.3 | 2307 | 21 | AAZ99938 |
| 17 | 98    | 9.8  | 2968 | 22 | AAH78342 |
| 18 | 96.2  | 9.6  | 1087 | 21 | AAZ99940 |
| 19 | 95.2  | 9.5  | 1036 | 21 | AAZ99939 |
| 20 | 95.2  | 9.5  | 2913 | 19 | AAV38719 |
| 21 | 94.8  | 9.5  | 1069 | 19 | AAV38721 |
| 22 | 93.6  | 9.4  | 1919 | 19 | AAV38722 |
| 23 | 93.6  | 9.4  | 3003 | 17 | AA42634  |
| 24 | 93.6  | 9.4  | 3090 | 19 | AAV38720 |
| 25 | 93    | 9.3  | 1120 | 21 | AAZ99941 |
| 26 | 92.8  | 9.3  | 2529 | 17 | AA42637  |
| 27 | 92.6  | 9.3  | 3074 | 18 | AA69587  |
| 28 | 92    | 9.2  | 2531 | 17 | AA417267 |
| 29 | 92    | 9.2  | 3033 | 17 | AA42630  |
| 30 | 90.4  | 9.0  | 2576 | 17 | AA42636  |
| 31 | 90.4  | 9.0  | 2578 | 17 | AA42631  |
| 32 | 90.4  | 9.0  | 2975 | 17 | AA42635  |
| 33 | 90.4  | 9.0  | 3231 | 17 | AA42632  |
| 34 | 80.8  | 8.1  | 2715 | 21 | AA645939 |
| 35 | 79.2  | 7.9  | 1452 | 21 | AA36957  |
| 36 | 67.4  | 6.7  | 2899 | 22 | AAH02926 |
| 37 | 67.4  | 6.7  | 3075 | 24 | AA594880 |
| 38 | 65.6  | 6.6  | 359  | 21 | AAA67263 |
| 39 | 63.6  | 6.4  | 408  | 21 | AAA67266 |
| 40 | 60.8  | 6.1  | 2687 | 20 | AA34646  |
| 41 | 60.8  | 6.1  | 4563 | 22 | AA30910  |
| 42 | 60.4  | 6.0  | 3128 | 16 | AA00774  |
| 43 | 59.2  | 5.9  | 2487 | 18 | AA69747  |
| 44 | 59.2  | 5.9  | 2565 | 18 | AA69752  |
| 45 | 59.2  | 5.9  | 2713 | 19 | AAV70962 |

## ALIGNMENTS

|          |  |                                  |
|----------|--|----------------------------------|
| RESULT 1 | AA34650  | AA34650 standard; DNA; 11473 BP. |
| ID       | AA34650  |                                  |
| AC       | AA34650;   |                                  |
| DT       | 05-JUL-1999  | (first entry)                    |
| XX       |  |                                  |
| XX       |  |                                  |
| DE       | Starch branching enzyme II (SBE II) gene sequence.                     |                                  |
| XX       |  |                                  |
| KW       | Starch biosynthetic pathway; cereal plant; enzyme; SBE; SSS; DBE; GBS; |                                  |
| KW       | starch branching enzyme; starch soluble synthase; debranching enzyme;  |                                  |
| KW       | endosperm; wheat; barley; granule-bound synthase; glutenin; starch;    |                                  |
| KW       | grain softness protein I; bacterial isoamylase; glycogen synthase;     |                                  |
| KW       | WSE I-D4 gene; ss.   |                                  |
| XX       |  |                                  |
| OS       | Triticum tauschii.   |                                  |
| XX       |  |                                  |
| PN       | W09914314-A1.  |                                  |
| XX       |  |                                  |
| PD       | 25-MAR-1999.   |                                  |
| XX       |  |                                  |
| PF       | 11-SEP-1998;   | 98WO-A000743.                    |
| XX       |  |                                  |
| PR       | 20-MAR-1998;   | 98AU-0002509.                    |
| PR       | 12-SEP-1997;   | 97AU-0009108.                    |
| XX       |  |                                  |
| PA       | (GSTR ) COMMONWEALTH SCI & IND RES ORG.                                |                                  |
| PA       | (GOOD-) GOODMAN FIELDER LTD.   |                                  |
| PA       | (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.                                 |                                  |
| PA       | (AUSU ) UNIV AUSTRALIAN NAT.   |                                  |
| XX       |  |                                  |

Corn starch branch  
Corn starch branch  
DNA encoding maize  
Plasmid pBE240 ins  
Zea mays starch br  
Rice starch branch  
DNA encoding part  
Nucleotide sequenc  
Nucleotide sequenc  
Nucleotide sequenc  
Full length cassav  
Class A starch bra  
Potato starch bran  
Class A starch bra  
Class A starch bra  
Class A starch bra  
Class A starch bra  
Arabidopsis thalia  
Arabidopsis thalia  
Human shear stress  
Human DNA sequenc  
Eucalyptus grandis  
Pinus radiata bran  
WSE I-D4 CDNA seq  
Wheat starch bran  
Potato starch bran  
Corn starch branch  
Corn starch branch  
DNA encoding maize

PI Li Z, Morell M, Rahman S;  
 XX  
 DR WPI, 1999-229525/19.  
 XX  
 PT New isolated cereal plant enzyme genes used for, e.g. expression of  
 PT antisense sequences of granule bound synthase  
 XX  
 XX  
 PS Claim 8; Page 75-81; 171pp; English.  
 XX  
 CC The invention relates to a novel enzyme of starch biosynthetic pathway  
 CC in a cereal plant, where the enzyme is selected from starch branching  
 CC enzyme (SBE) I, SBE II, starch soluble synthase (SSS) I, and debranching  
 CC enzyme (DBE), with the proviso that the enzyme is not SSS I of rice, or  
 CC SBE I of rice or maize. The methods and products can be used for  
 CC targeting expression specifically to the endosperm of the seeds of cereal  
 CC plants such as wheat or barley. They can be used for the expression of  
 CC e.g. antisense sequences of granule-bound synthase (GBSS), SBE II, low  
 CC mol. wt. glutenin, grain softness protein I, bacterial isomylase,  
 CC bacterial glycogen synthase, and wheat high mol. wt. glutenin Bx17. They  
 CC can be used for modifying the characteristics of starch produced by a  
 CC plant. The present sequence represents the SBE II gene sequence.  
 XX  
 SQ Sequence 11473 BP; 3096 A; 2263 C; 2423 G; 3691 T; 0 other;

|                            |         |                     |        |                   |
|----------------------------|---------|---------------------|--------|-------------------|
| Query Match                | 100.0%; | Score 1001;         | DB 20; | Length 11473;     |
| Best Local Similarity      | 100.0%; | Pred. No. 3.3e-249; |        |                   |
| Matches 1001; Conservative | 0;      | Mismatches          | 0;     | Indels 0; Gaps 0; |

|    |       |  |       |
|----|-------|--|-------|
| OY | 1     | ttggcagatgaacagatgaatatataagcctctcttcttgagagatatgatattca           | 60    |
| Db | 9510  | tcgcagatagaacagatgaatatataagcctctcttcttgagagatatgatattca           | 9569  |
| OY | 61    | tggctctgatatagcttcaactcttcgcacttgatcgttgcgatagcattacataaaatgat     | 120   |
| Db | 9570  | tggctctgatatagcttcaactcttcgcacttgatcgttgcgatagcattacataaaatgat     | 9629  |
| OY | 121   | caggctcttcaccaaagggtttatggtggtgaaggtctatcttaacttcaatgagaaatgagtt   | 180   |
| Db | 9630  | caggctcttcaccaaagggtttatggtggtgaaggtctatcttaacttcaatgagaaatgagtt   | 9689  |
| OY | 181   | tgggcatcctgttcaagctcttcaacaacattatggcattctgcgcatgatgtgatattactgta  | 240   |
| Db | 9690  | tgggcatcctgttcaagctcttcaacaacattatggcattctgcgcatgatgtgatattactgta  | 9749  |
| OY | 241   | atttgaacaaatgcttctcttccacattgataatgataatgtaactcgttgccttccaaagag    | 300   |
| Db | 9750  | atttgaacaaatgcttctcttccacattgataatgataatgtaactcgttgccttccaaagag    | 9809  |
| OY | 301   | gaagttaaacttattattacttggcagaatggatagatttccaaaggcgcccaaacctc        | 360   |
| Db | 9810  | gaagttaaacttattattacttggcagaatggatagatttccaaaggcgcccaaacctc        | 9869  |
| OY | 361   | tccaaacccggcaaaagtctccctctgggaataaacaataagatgataataaattgcgcgctagat | 420   |
| Db | 9870  | tccaaacccggcaaaagtctccctctgggaataaacaataagatgataataaattgcgcgctagat | 9929  |
| OY | 421   | ttgatcttgttaagtttttagctgtgtcattatacatctccctcaccagatctttatttggccatt | 480   |
| Db | 9930  | ttgatcttgttaagtttttagctgtgtcattatacatctccctcaccagatctttatttggccatt | 9989  |
| OY | 481   | tattctctgataaatacaatacgtttgtgttagaagaagatacaacatgcctttttagcttgc    | 540   |
| Db | 9990  | tattctctgataaatacaatacgtttgtgttagaagaagatacaacatgcctttttagcttgc    | 10049 |
| OY | 541   | tagacgttaacaataagtatgtgtgcagagttgtgatacatataaanaatcatgatattttt     | 600   |
| Db | 10050 | tagacgttaacaataagtatgtgtgcagagttgtgatacatataaanaatcatgatattttt     | 10109 |
| OY | 601   | gcaaggagatgagaattttcttaagatatcogtattgtgcaaggttcogatacgaagcaatga    | 660   |
| Db | 10110 | gcaaggagatgagaattttcttaagatatcogtattgtgcaaggttcogatacgaagcaatga    | 10169 |

|    |       |  |       |
|----|-------|--|-------|
| QY | 661   | gctctttaggaaaaaatatggggatgtgcacgcggttgccttgcgtacaaagtcac         | 720   |
| Db | 10170 | gcactcttaggaaaaatatgtggatgtgcacgcggttgccttgcgtacaaagtcac         | 10229 |
| QY | 721   | agttacgcgtcagtcctcttcaagtcggtgtaaaaaagtgtagaatctccgttaatgaga     | 780   |
| Db | 10230 | agttacgcgtcagtcctcttcaagtcggtgtaaaaaagtgtagaatctccgttaatgaga     | 10289 |
| QY | 781   | tgaaacacgtgcgaaggcgagcgtgaattgtcttccaccaaactattcttctaagtgc       | 840   |
| Db | 10290 | tgaaacacgtgcgaaggcgagcgtgaattgtcttccaccaaactattcttctaagtgc       | 10349 |
| QY | 841   | ttgtgtattgtatacatataccagcacctgacgaattgaactgcagttatgaaactctgagga  | 900   |
| Db | 10350 | ttgtgtattgtatacatataccagcacctgacgaattgaactgcagttatgaaactctgagga  | 10409 |
| QY | 901   | ccagatgctgttccacggaacaacatgaggaagataagtgatcatcaccctcaaaaggagattt | 960   |
| Db | 10410 | ccagatgctgttccacggaacaacatgaggaagataagtgatcatcaccctcaaaaggagattt | 10469 |
| QY | 961   | ggatattgttttccaacttccactctgagacatgcttttttg                       | 1001  |
| Db | 10470 | ggatattgttttccaacttccactctgagacatgcttttttg                       | 10510 |

| RESULT   | 2                                  |
|----------|------------------------------------|
| AAH78338 |                                    |
| ID       | AAH78338 standard; cDNA; 11475 BP. |

DT 26-NOV-2001 (first entry)

DE Nucleotide sequence of a starch branching enzyme designated F2.

Wheat; starch branching enzyme; BEIIb; SBE; transgenic plant;  
starch biosynthetic pathway; amylopectin; F2; amylose; ss.

OS *Aegilops tauschii*.

PN WO200162934-A1.

PD 30-AUG-2001.

PF 21-FEB-2001; 2001WO-AU00175.

PR 21-FEB-2000; 2000AU-0005742.

PA (CSIR ) COMMONWEALTH SCI &amp; IND RES ORG.

PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

PI Morell M, Rahman S;

DR WPI; 2001-570635/64.

PT Nucleic acids encoding wheat starch branching enzyme IIb, useful for

PT wheat and barley -

PS Example 1; Fig 2; 103pp; English.

CC The present sequence encodes a wheat starch branching enzyme of

CC ancestral D genome donor of wheat. Probes isolated from the present

CC in wheat, especially BEIb. The BEIb nucleic acids may be used to

CC altering their nutritional content by modulating the starch

CC produced in the plant.

**SQ** Sequence 11475 BP; 3093 A; 2259 C; 2423 G; 3689 T; 11 other,

Query Match 97.7% Score 978; DB 22; Length 11475;  
 Best Local Similarity 99.8%; Pred. No. 3e-243;  
 Matches 1000; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 1 tccgcacatagacacgacgaacatcaacgctctcttctgttcagagatagtatgttca 60  
 |||||||  
 DB 9512 tccgcacatagacacgacgaacatcaacgctctcttctgttcagagatagtatgttca 9571

OY 61 tgcctcttgatagag-cttcaactcttcgcacatgcgtgcgacatagacatataatga 119  
 |||||||  
 DB 9572 tgcctcttgatagagccttcaactcttcgcacatgcgtgcgacatagacatataatga 9631

OY 120 tcaagctgttcaacacatgggtttaggttggtgaagcgtacatcctaacttcagtggaatgagt 179  
 |||||||  
 DB 9632 tcaagctgttcaacacatgggtttaggttggtgaagcgtacatcctaacttcagtggaatgagt 9691

OY 180 ttgggcacctcgttcagctcttcaacacatatttgcatcttcgacatgttgatcttactgt 239  
 |||||||  
 DB 9692 ttgggcacctcgttcagctcttcaacacatatttgcatcttcgacatgttgatcttactgt 9751

OY 240 aattgaaacacatgctcttctcttccacatgtatgtatctgaatcgttgcctccaagga 299  
 |||||||  
 DB 9752 aattgaaacacatgctcttctcttccacatgtatgtatctgaatcgttgcctccaagga 9811

OY 300 ggaagttactcttacttactcttgcagaaatggatagatttcccaagagcccaaacac 359  
 |||||||  
 DB 9812 ggaagttactcttacttactcttgcagaaatggatagatttcccaagagcccaaacac 9871

OY 360 ttccacacggcgaagattctccctcgtgaataacaaatagttatgataaagtccgcgctaga 419  
 |||||||  
 DB 9872 ttccacacggcgaagattct-ccctcgtgaataacaaatagttatgataaagtccgcgctaga 9930

OY 420 ttgcatcttgtaagtttagctgtgcatactacatccctcactagacatcttattggccat 479  
 |||||||  
 DB 9931 ttgcatcttgtaagtttagctgtgcatactacatccctcactagacatcttattggccat 9990

OY 480 ttattcttgatgaatacataatgtttttaggaaagatcaacatgctcttggtagtttt 539  
 |||||||  
 DB 9991 ttattcttgatgaatacataatgtttttaggaaagatcaacatgctcttggtagtttt 10050

OY 540 gtagacgttaacataagatgtgttgagagttgttgatcatatnaaaatatacagttttt 599  
 |||||||  
 DB 10051 gtagacgttaacataagatgtgttgagagttgttgatcatatnaaaatatacagttttt 10110

OY 600 tgcagggagatgcagatttctcttagatacgtgtgcatacgaagatcgcatacggcaatgc 659  
 |||||||  
 DB 10111 tgcagggagatgcagatttctcttagatacgtgtgcatacgaagatcgcatacggcaatgc 10170

OY 660 agcatcttgaggaataaatacagtggtatcactgtgttgccttctgttcataacaagtca 719  
 |||||||  
 DB 10171 agcatcttgaggaataaatacagtggtatcactgtgttgccttctgttcataacaagtca 10230

OY 720 cagtttaacgtcagctctctcacaagtgtaaaaaaagttagaataatcccgtaatatgag 779  
 |||||||  
 DB 10231 cagtttaacgtcagctctctcacaagtgtaaaaaaagttagaataatcccgtaatatgag 10290

OY 780 atgaaactgtgcagaaagcgagctggaattgctttccacccaacatatttcttaagt 839  
 |||||||  
 DB 10291 atgaaactgtgcagaaagcgagctggaattgctttccacccaacatatttcttaagt 10350

OY 840 ctgtgtatttgatatacatcacagcactgacaaatgtaactgtaagttagacatctgagc 899  
 |||||||  
 DB 10351 ctgtgtatttgatatacatcacagcactgacaaatgtaactgtaagttagacatctgagc 10410

OY 900 accaatatgtttccacgaaacatgaggaagataaagttgatacctcctcaaaagaggagatt 959  
 |||||||  
 DB 10411 accaatatgtttccacgaaacatgaggaagataaagttgatacctcctcaaaagaggagatt 10470

OY 960 tggatatgttttcaacttcacatgaggaacatagcttttttg 1001  
 |||||||  
 DB 10471 tggatatgttttcaacttcacatgaggaacatagcttttttg 10512

RESULT 3  
 AAZ35393  
 ID AAZ35393 standard; DNA; 23449 BP.  
 XX  
 AC AAZ35393;  
 XX  
 DT 27-MAR-2000 (first entry)  
 XX  
 DE Maize amylose-extender (Ae) gene encoding starch branching enzyme.  
 XX  
 KW Starch branching enzyme IIB; amylose-extender gene; Ae gene; maize;  
 KW corn; promoter; expression control element; endosperm;  
 KW transgenic plant; ss.  
 XX  
 OS Zea mays.  
 XX  
 FH  
 FT Key' location/Qualifiers  
 FT misc\_signal 597..602  
 FT /tag= a  
 FT /note= "I box, conserved in RBCS genes"  
 FT repeat\_unit 1475..1481  
 FT /tag= b  
 FT /note= "RY repeat, distal regulatory element"  
 FT misc\_signal 1967..1973  
 FT /tag= c  
 FT /note= "Hex, conserved element found in plant  
 FT histone gene promoters"  
 FT misc\_signal 2593..2599  
 FT /tag= d  
 FT /note= "MRE box, motif essential for metal  
 FT ion-dependent induction of metallothionein  
 FT genes"  
 FT GC\_signal 2666..2671  
 FT /tag= e  
 FT GC\_signal 2738..2743  
 FT /tag= f  
 FT GC\_signal 2831..2836  
 FT /tag= g  
 FT misc\_signal 2838..2844  
 FT /tag= h  
 FT /note= "MRE box, motif essential for metal  
 FT ion-dependent induction of metallothionein  
 FT genes"  
 FT GC\_signal 2874..2879  
 FT /tag= i  
 FT TATA\_signal 2933..2936  
 FT /tag= j  
 FT CDS 3065..19659  
 FT /tag= k  
 FT /note= "contains introns"  
 FT exon 2965..3176  
 FT /tag= l  
 FT intron 3177..3282  
 FT /tag= m  
 FT exon 3283..3428  
 FT /tag= n  
 FT intron 3429..3672  
 FT /tag= o  
 FT exon 3673..3827  
 FT /tag= p  
 FT intron 3828..4913  
 FT /tag= q  
 FT exon 4914..5012  
 FT /tag= r  
 FT intron 5013..5112  
 FT /tag= s  
 FT exon 5113..5212  
 FT /tag= t  
 FT intron 5213..5312  
 FT /tag= u  
 FT exon 5313..5412  
 FT /tag= v  
 FT intron 5413..5512  
 FT /tag= w  
 FT exon 5513..5612  
 FT /tag= x  
 FT intron 5613..5712  
 FT /tag= y  
 FT exon 5713..5812  
 FT /tag= z  
 FT intron 5813..5912  
 FT /tag= 1  
 FT exon 5913..6012  
 FT /tag= 2  
 FT intron 6013..6112  
 FT /tag= 3  
 FT exon 6113..6212  
 FT /tag= 4



```

Db 17932 gtttaaggagagggctatcttaattcatggaagatgtttgacatcctgtgaga 17991
QY 198 ctttaacaattatgcatcttcgcatgatgtgattactgta-----ttgaac 248
Db 17992 tttaactacttgcttcatcttaacccctgcgtgagctcttataagacagtaacctcaac 18051
QY 249 catgctctcttccatcatgtaatgtaattatgtaactgtgtgctccaagaggaattaa 308
Db 18052 aattactctgcaattatctcttgttagttatataatgtagttagtgagcttgagatcttt 18111
QY 309 ctctctatcttgatgagagaatgatagtatctccaagaagccacaactctccaacg 368
Db 18112 ctctctatcttatttgcaagaatgatagtatctccaagaaggtccgcaagaagcttcaagt 18171
QY 369 gcaaaagtctcccttggaataaataagtaagtaagataagtcgcgcgctgattgattctt 428
Db 18172 gt-aagttatctccagggagataacacacagctaaagacaaagtcgcgcgaagattgacgt 18230
QY 429 gtaagtttagctgtgctatatactccctcaactagatacttattgcca-----478
Db 18231 gtaactctcttgattgctgcaaaagtcgaagttgatttacttaccacatgcacg 18290
QY 479 -----ttattcttgatgaataatagtgtg 507
Db 18291 tgccttaacgaagagtgccatataaataagttccctcttattcatatagatgttcc 18350
QY 508 ttggaagaatcaacatgtgcttgtt-aggtttgaagcgttaagataagtagtgtg 565
Db 18351 agaagaataaatacttactactctgttaaaagtttcttataatagtcgcttgttgg 18410
QY 566 agagttgttgcataataaataatcatgattttt-----gcaaggagatcaagatt 618
Db 18411 tcaataatcgcatcatctcggaattgttatttatttaccatgtgcaggtgagtagcagacta 18470
QY 619 tcttaataatcgtggtatgcaagagttcgatcagcgaacagcagcatcttgaggaataa 678
Db 18471 tcttaagtaatacgtgatacgaagtttgcagcgaatgcacacatcttgagcaaaata 18530
QY 679 tggggtatg-----tcactgtgtgtcttgttgatgataaagtcacagttt 725
Db 18531 tgaagatgtctcttcttacttcttacttcttgcattgtcttcgaagttccacaacatcat 18590
QY 726 aaagtcagctcttccaagtgtgtaaaaaaagtgtagaattaatccctg-----taatg 777
Db 18591 attgtgtgcatcttactactgtaatgtaatctttaaanaaaatccatccctcagtttactg 18650
QY 778 agatgaaaactgtgcaaaagcgagctgtaattgct-----ttcaccaaaactatttc 832
Db 18651 agcttttaagcaatggaagtttcatatgaattcttcatatgtaatacactcttag 18710
QY 833 ttaagtgcctgtgtattgata-----883
Db 18711 gtaatttcaagatcatataaagtaactctgagacagcagcaataatgtaacgaaaaa 18770
QY 854 catataccagcagctgcaaatgtactgcagttatgacatctgagacagcagtaatttca 913
Db 18771 ttcttgatgaataatttgcgtgtaattgacagttcatgaacatctgataccagatatttcc 18830
QY 914 cggaaacaatgaggaataagatgatacctccaaagaggaattgtgattgttttc 973
Db 18831 cggaaacaatgaggaataagatgatacctccaaagaggaattgtgattgttttc 18890
QY 974 aacttcactgagcaatagctttttg 1001
Db 18891 aacttcactgcaacaacagcatatttg 18918

```

```

RESULT 4
AAH78337
ID AAH78337 standard; cDNA; 2726 BP.
XX
AC AAH78337;
XX

```

```

DT 26-NOV-2001 (first entry)
XX
DE Nucleotide sequence of wheat starch branching enzyme 9 (BEIIa).
XX
KW Wheat; starch branching enzyme; BEIIa; BEIIb; SBE; transgenic plant;
KW starch biosynthetic pathway; amylopectin; amylose; ss.
XX
OS Triticum sp.
XX
PN WO200162934-A1.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001MO-AU00175.
XX
PR 21-FEB-2000; 2000AU-0005742.
XX
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA (GOOD-) GOODMAN FIELDER LTD.
PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.
XX
PI Morell M, Rahman S;
XX
DR WPI; 2001-570635/64.
XX
PT Nucleic acids encoding wheat starch branching enzyme IIb, useful for
PT altering the amylose and amylopectin content of cereal plants, e.g.
PT wheat and barley -
XX
PS Example 1; Fig 1; 103pp; English.
XX
CC The present sequence encodes wheat starch branching enzyme 9,
CC designated BEIIa. The specification describes BEIIb. BEIIb is a
CC type II starch branching enzyme (SBE). The BEIIb nucleic acids may
CC be used to genetically transform cereal plants such as wheat or
CC barley and for altering their nutritional content by modulating
CC the starch biosynthetic pathway to vary levels of amylopectin and/or
CC amylose produced in the plant.
XX
SQ Sequence 2726 BP; 726 A; 564 C; 701 G; 735 T; 0 other;

```

```

Query Match 13.6%; Score 136.4; DB 22; Length 2726;
Best Local Similarity 98.7%; Pred. No. 1.9e-25;
Matches 148; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 43 aggaatgatgatcttcaatgctctgtagatag-cttcaactcttcgcatgtatcgga 101
Db 1787 aggaatgatgatcttcaatgctctgtagatagcttcaactcttcgcatgtatcgga 1846
QY 102 tagcatcataaataagatcagctgtgtaaccatgggttagtgtggaagctatctta 161
Db 1847 tagcatcataaataagatcagctgtgtaaccatgggttagtgtggaagctatctta 1906
QY 162 acttcattgggaatgagttgggacatcgt 191
Db 1907 acttcattgggaatgagttgggacatcgt 1936

```

```

RESULT 5
AAZ99942
ID AAZ99942 standard; DNA; 979 BP.
XX
AC AAZ99942;
XX
DT 25-JUL-2000 (first entry)
XX
DE Nucleotide sequence of starch branching enzyme II clone B1.
XX
KW Wheat; starch branching enzyme II; SBEII; SBEII-2; starch; SBEII-1;
KW gelatinisation onset; transgenic plant; foodstuff; bakery product; ss.
XX
OS Triticum aestivum.

```

|    |  |  |
|----|--|--|
| XX | WM0200015810-AL.   |  |
| XX | 23-MAR-2000.   |  |
| XX | 09-SEP-1999;   | 99WO-GB03011.                          |
| XX | 10-SEP-1998;   | 98EP-0307337.                          |
| XX | (PLAN-) PLANT BREEDING INT CAMBRIDGE LTD.                                |  |
| XX | Goldsbrough A, COLLIVER S;   |  |
| XX | WPI; 2000-271446/23.   |  |
| XX | DNA encoding wheat starch branching enzyme II isoforms, useful in        |  |
| XX | altering the characteristics of a plant, especially elevated starch      |  |
| XX | gelatinization onset and/or peak temperature -                           |  |
| XX | Claim 5; Page 124; 197pp; English.                                       |  |
| XX | The present sequence represents a clone of wheat starch branching        |  |
| XX | enzyme II (SBEII), which is part of a novel subclass of known SBEII      |  |
| XX | genes, designated SBEII-2, subclass B (SBEII-2B). The specification also |  |
| XX | describes subclass SBEII-1 genes, which are thought to have similar      |  |
| XX | functional properties to the maize SBEIIb gene. Starch branching enzymes |  |
| XX | catalyse the formation of the alpha-1,6 linkages, creating branch points |  |
| XX | in the growing starch molecule, via hydrolysis of an alpha-1,4 linkage   |  |
| XX | followed by reattachment of the released alpha-1,4-glucan chain to the   |  |
| XX | same or another glucosyl chain. SBEII polypeptides can be used to alter  |  |
| XX | the characteristics of a plant, in particular to alter starch so that    |  |
| XX | it has an elevated gelatinisation onset and/or peak temperature. Starch  |  |
| XX | obtained from transgenic plants is useful in the preparation or          |  |
| XX | processing a foodstuff, particularly bakery products.                    |  |
| XX | Sequence 979 BP; 271 A; 203 C; 247 G; 257 T; 1 other;                    |  |
| XX |  |  |
| XX | Query Match  | 13.3%; Score 133.4; DB 21; Length 979; |
| XX | Best Local Similarity  | 98.6%; Pred. No. 8.4e-25;              |
| XX | Matches 145; Conservative  | 0; Mismatches 1; Indels 1; Gaps 1,     |
| XX |  |  |
| XX | 46 atatgatattcttcgtctgctgtagg-cttcaactcttcgattgactgtgcatg 104            |  |
| XX | 1 atatgtatgattcttaagctcgcgataaggcttcaactcttcgattgactgtgcatg 60           |  |
| XX | 105 cattacataaaatgatacagctctgcacacatggttagtggatgaagctatcttaact 164       |  |
| XX | 61 cattacataaaatgatacagctctgcacacatggttagtggatgaagctatcttaact 120        |  |
| XX | 165 tcatgggaatgagtttgggcatcctg 191                                       |  |
| XX | 121 tcatgggaatgagtttgggcatcctg 147                                       |  |
| XX |  |  |
| XX | RESULT 6   |  |
| XX | AAZ99947   |  |
| XX | AAZ99947 standard; DNA; 977 BP.  |  |
| XX | AAZ99947;  |  |
| XX | 25-JUL-2000 (first entry)  |  |
| XX |  |  |
| XX | Nucleotide sequence of starch branching enzyme II clone B11.             |  |
| XX |  |  |
| XX | Wheat; starch branching enzyme II; SBEII; SBEII-1; starch; SBEII-2;      |  |
| XX | gelatinisation onset; transgenic plant; foodstuff; bakery product; ss.   |  |
| XX | Triticum aestivum.   |  |
| XX |  |  |
| XX | Key  | Location/Qualifiers                    |
| XX | FT   | 3..641                                 |
| XX | FT   | /*tag- a                               |

|                           |  |  |
|---------------------------|--|--|
| FT                        |  | /product= "starch branching enzyme II" |
| XX                        |  |  |
| PN                        | WO200015810-A1.  |  |
| XX                        |  |  |
| PD                        | 23-MAR-2000.   |  |
| XX                        |  |  |
| PF                        | 09-SEP-1999;   | 99WO-CB03011.                          |
| XX                        |  |  |
| PX                        | 10-SEP-1998;   | 98EP-0307337.                          |
| XX                        |  |  |
| PA                        | (PLAN-) PLANT BREEDING INT CAMBRIDGE LTD.                                |  |
| XX                        |  |  |
| PI                        | Goldsbrough A, Collier S;  |  |
| DR                        | WPI: 2000-271446/23.   |  |
| XX                        | P-FSDB; AAU84411.  |  |
| PT                        | DNA encoding wheat starch branching enzyme II isoforms, useful in        |  |
| PR                        | altering the characteristics of a plant, especially elevated starch      |  |
| PS                        | gelatinization onset and/or peak temperature -                           |  |
| XX                        |  |  |
| PS                        | Disclosure; Page 168-169; 197pp; English.                                |  |
| XX                        |  |  |
| CC                        | The present sequence represents a clone of wheat starch branching        |  |
| CC                        | enzyme II (SBEII), which is part of a novel subclass of known SBEII      |  |
| CC                        | genes, designated SBEII-1. The specification also describes subclasses   |  |
| CC                        | SBEII-2 sequences. The SBEII-1 genes are thought to have similar         |  |
| CC                        | functional properties to the maize SBEIIb gene. Starch branching enzymes |  |
| CC                        | catalyse the formation of the alpha-1,6 linkages, creating branch points |  |
| CC                        | in the growing starch molecule, via hydrolysis of an alpha-1,4 linkage   |  |
| CC                        | followed by reattachment of the released alpha-1,4-glucan chain to the   |  |
| CC                        | same or another glucosyl chain. SBEII polypeptides can be used to alter  |  |
| CC                        | the characteristics of a plant, in particular to alter starch so that    |  |
| CC                        | it has an elevated gelatinisation onset and/or peak temperature. Starch  |  |
| CC                        | obtained from transgenic plants is useful in the preparation or          |  |
| CC                        | processing a foodstuff, particularly bakery products.                    |  |
| SQ                        | Sequence 977 BP; 270 A; 200 C; 245 G; 262 T; 0 other;                    |  |
| Query Match               | 13.0%; Score 130.2; DB 21; Length 977;                                   |  |
| Best Local Similarity     | 97.3%; Pred. No. 5.7e-24;  |  |
| Matches 143; Conservative | 0; Mismatches 3; Indels 1; Gaps  | 1.                                     |
| OY                        | 46 atatgtatgatttcacgtgccttgtagag-ggtccaactcttcgatcgttgcgtgcataag 104     |  |
| Db                        | 1 ataagtatgatttcacgtgccttgtagagccctcaactcttcgatcgttgcgtgcataag 60        |  |
| OY                        | 105 cattcataaaatgtagcggtctgcacatgggtttggttggaaggctatctaact 164           |  |
| Db                        | 61 cattcataaaatgtagcgaagcttgcacatgggtttggttggaaggctatctaact 120          |  |
| OY                        | 165 tcatggaaatgagttggcacctcty 191  |  |
| Db                        | 121 tcatggaaatgagttggcacctcty 147  |  |
| RESULT 7                  |  |  |
| ID                        | AAZ99946   |  |
| XX                        | AAZ99946 standard; DNA; 984 BP.  |  |
| AC                        | AAZ99946;  |  |
| DT                        | 25-JUL-2000 (first entry)  |  |
| DE                        | Nucleotide sequence of starch branching enzyme II clone A2.              |  |
| KM                        | Wheat; starch branching enzyme II; SBEII; SBEII-1; starch; SBEII-2;      |  |
| XX                        | gelatinisation onset; transgenic plant; foodstuff; bakery product; ss.   |  |
| OS                        | Triticum aestivum.   |  |
| PN                        | WO200015810-A1.  |  |





DR WPI; 2000-271446/23.

XX DNA encoding wheat starch branching enzyme II isoforms, useful in  
 PT altering the characteristics of a plant, especially elevated starch  
 PT gelatinization onset and/or peak temperature -  
 XX  
 XX Disclosure; Page 187; 197pp; English.

XX The present sequence represents the consensus sequence of wheat starch  
 CC branching enzyme II (SBEII). The specification describes novel subclasses  
 CC of SBEII, designated SBEII-1 and SBEII-2. The SBEII-1 genes are thought  
 CC to have similar functional properties to the maize SBEIIb gene. Starch  
 CC branching enzymes catalyze the formation of the alpha-1,6 linkages,  
 CC creating branch points in the growing starch molecule, via hydrolysis of  
 CC an alpha-1,4 linkage followed by reattachment of the released alpha-1,4-  
 CC glucan chain to the same or another glucosyl chain. SBEII polypeptides  
 CC can be used to alter the characteristics of a plant, in particular to  
 CC alter starch so that it has an elevated gelatinization onset and/or peak  
 CC temperature. Starch obtained from transgenic plants is useful in the  
 CC preparation or processing a foodstuff, particularly bakery products.

SQ Sequence 1085 BP; 268 A; 185 C; 284 G; 348 T; 0 other:

Query Match 11.4%; Score 114.2; DB 21; Length 1085;  
 Best Local Similarity 90.5%; Pred. No. 8.2e-20;  
 Matches 133; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

OY 46 atatgataatttcattcgtctgtgatatg-gcttcaactcttcgtatggtgagatg 104  
 Db 1 atatgataatttcattcgtctgtgatatg-gcttcaactcttcgtatggtgagatg 60  
 OY 105 cattacataaattgatcaggctgttcacacatggtttagtggtagaaggctacttaact 164  
 Db 61 cattgacataaattgatgagctgttcacacatggtttagtggtagaaggctacttaact 120  
 OY 165 tcatgggaaatgagttgggcatctcg 191  
 Db 121 tcatgggaaatgagttgggcatctcg 147

RESULT 10  
 AAT69737  
 ID AAT69737 standard; DNA; 2087 BP.  
 XX  
 AC AAT69737;  
 XX  
 DT 10-SEP-1997 (first entry)  
 XX  
 DE Corn starch branching enzyme IIB gene fragment in pBE96.  
 XX  
 KW Starch branching enzyme IIB; SBEIIB; corn; maize; antisense;  
 KW amylopectin; transgenic plant; pBE96; ss.  
 XX  
 OS Zea mays.  
 XX  
 PN WO9722703-A2.  
 XX  
 PD 26-JUN-1997.  
 XX  
 PF 12-DEC-1996; 96WO-US19678.  
 XX  
 PR 20-DEC-1995; 95US-0009113.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E. I.  
 XX  
 PI Broglie KE, Hubbard NL, Klein TW;  
 XX  
 DR WPI; 1997-341694/31.  
 XX  
 PT Transgenic corn in which grain derived starch fine structure is  
 PT controlled - specifically branch chain distribution of amylopectin,  
 PT useful in preparation of thickened foodstuff

XX Example 4; Page 56-57; 92pp; English.

PS A cDNA insert (AAT69737) in plasmid pBE96 comprises a 2.09 kb  
 XX fragment of corn starch branching enzyme IIB (SBEIIB) cDNA (see  
 CC also AAT69729), starting at the initiating ATG codon of the coding  
 CC region and terminating 312 bp 5' of the translation termination  
 CC codon. It was obt. by site-specific mutagenesis of the SBEIIB  
 CC gene in plasmid pBE240 to generate an NcoI site at the ATG start  
 CC site, and EcoRI digestion of the mutagenized plasmid. The 2.09 kb  
 CC fragment was inserted, in sense orientation, between the zein  
 CC promoter and 3' regions in vector pM103 to produce pBE96.  
 CC Co-suppression and antisense techniques can be used to inhibit  
 CC SBEIIB in transgenic plants and thereby alter starch properties.

SQ Sequence 2087 BP; 565 A; 396 C; 553 G; 573 T; 0 other:

Query Match 11.1%; Score 110.8; DB 18; Length 2087;  
 Best Local Similarity 88.0%; Pred. No. 7.6e-19;  
 Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

OY 43 aggatatgatatttcattcgtctgtgatatg-gcttcaactcttcgtatggtgagca 101  
 Db 1757 aggatatgatatttcattcgtctgtgatatg-gcttcaactcttcgtatggtgagca 1816  
 OY 102 tagcattacataaattgatcaggctgttcacacatggtttagtggtagaaggctactta 161  
 Db 1817 tagcattacataaattgatcaggctgttcacacatggtttagtggtagaaggctactta 1876  
 OY 162 acttcattgggaaatgagttgggcatctcg 191  
 Db 1877 acttcattgggaaatgagttgggcatctcg 1906

RESULT 11  
 AAT69736/C  
 ID AAT69736 standard; DNA; 2165 BP.  
 XX  
 AC AAT69736;  
 XX  
 DT 10-SEP-1997 (first entry)  
 XX  
 DE Corn starch branching enzyme IIB gene fragment in pBE45.  
 XX  
 KW Starch branching enzyme IIB; SBEIIB; corn; maize; antisense;  
 KW amylopectin; transgenic plant; pBE45; ss.  
 XX  
 OS Zea mays.  
 XX  
 PN WO9722703-A2.  
 XX  
 PD 26-JUN-1997.  
 XX  
 PF 12-DEC-1996; 96WO-US19678.  
 XX  
 PR 20-DEC-1995; 95US-0009113.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E. I.  
 XX  
 PI Broglie KE, Hubbard NL, Klein TW;  
 XX  
 DR WPI; 1997-341694/31.  
 XX  
 PT Transgenic corn in which grain derived starch fine structure is  
 PT controlled - specifically branch chain distribution of amylopectin,  
 PT useful in preparation of thickened foodstuff

Example 3; Page 55-56; 92pp; English.

CC A 2.16 kb cDNA insert (AAT69736) in plasmid pBE45 comprises the  
 CC entire 5' untranslated region as well as 2.08 kb of the coding  
 CC region of corn starch branching enzyme IIB (SBEIIB) cDNA (see

CC also AAT69729). It was generated by EcoRI digestion of plasmid  
 CC pBE240 and was inserted into vector pMD103 in antisense orientation  
 CC with respect to a maize 27 kDa zein promoter and a 10 kDa zein gene  
 CC 3' sequence. Transgenic corn plants expressing the near full-  
 CC length SBEIIb antisense transcript demonstrated a unique starch  
 CC branching phenotype.

SQ Sequence 2165 BP; 593 A; 580 C; 416 G; 576 T; 0 other;

Query Match 11.1%; Score 110.8; DB 18; Length 2165;  
 Best Local Similarity 88.0%; Pred. No. 7.7e-19;  
 Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 43 aggatatgctatgtatcttcctgctgtag-gcttaactcttcgattgctgca 101  
 |||  
 DB 331 AGGATATGATGATTCATGCGCCCTCGATGACCTTCACTCCATCCATGATCGTGGGA 272  
 QY 102 tagcattacataaataatgataagcttgcacatgggttaggtgtggaagctatctta 161  
 |||  
 DB 271 TAGCATTACATAGATGATGACTTATCACAATGGGTTTAGAGAGAGAGGCTATCTTA 212  
 QY 162 acttcattgggaatgagttgggacatctg 191  
 |||  
 DB 211 ATTTCATGGCAATGAGTTGACATCCTG 182

RESULT 12  
 AAT69729

ID AAT69729 standard; DNA; 2640 BP.

AC AAT69729;

DT 23-AUG-1999 (first entry)

DE DNA encoding maize branching enzyme II.

DE Non-glycogen-like polysaccharide production; fermentation;

KW starch synthesis enzyme; glycogen-synthesis enzyme; glycogen synthesis;

KW non-starch branching gene; amylopectin; amylose; plant-like starch;

KW maize branching enzyme II; ss.

OS Zea mays.

PN MO9844780-A1.

PD 15-OCT-1998.

PF 03-APR-1998; 98WO-US06660.

PR 04-APR-1997; 97US-0042939.

PA (EXSE-) EXSEED GENETICS LLC.

PI Guan H, Keeling PJ;

DR WPI: 1998-568285/48.

DR P-PADB; AAT69729.

PT Producing non-glycogen-like polysaccharides in bacteria, fungi or

PT plants - transformed with genes for enzymes involved in starch or

PT glycogen synthesis allows fermentative production of starches with

PT engineered properties

PS Disclosure; Fig 53; 150pp; English.

XX The specification describes a method for the production of  
 CC non-glycogen-like polysaccharides in a host. The method comprises  
 CC transforming a host, suitable for fermentation, with genes encoding  
 CC starch- or glycogen-synthesis enzymes, and fermenting the transformants.  
 CC The specification also describes hosts transformed with a gene active  
 CC in glycogen synthesis and at least one non-starch branching gene,  
 CC involved in production of amylopectin or amylose in its original host.

CC The method is used to produce plant-like starches by fermentation and  
 CC new starches in plants. These starches are useful for all food and  
 CC non-food applications of starch. The present sequence is used in  
 CC the course of the invention.

SQ Sequence 2640 BP; 715 A; 510 C; 686 G; 729 T; 0 other;

Query Match 11.1%; Score 110.8; DB 19; Length 2640;  
 Best Local Similarity 88.0%; Pred. No. 8.2e-19;  
 Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 43 aggatatgctatgtatcttcctgctgtag-gcttaactcttcgattgctgca 101  
 |||  
 DB 1762 aggatatgctatgtatcttcctgctgtag-gcttaactcttcgattgctgca 1821  
 QY 102 tagcattacataaataatgataagcttgcacatgggttaggtgtggaagctatctta 161  
 |||  
 DB 1822 tagcattacataaataatgataagcttgcacatgggttaggtgtggaagctatctta 1881  
 QY 162 acttcattgggaatgagttgggacatctg 191  
 |||  
 DB 1882 atttcattgggaatgagttgggacatctg 1911

RESULT 13  
 AAT69729

ID AAT69729 standard; DNA; 2665 BP.

AC AAT69729;

DT 10-SEP-1997 (first entry)

DE Plasmid pBE240 insert encoding corn starch branching enzyme IIB.

KW Starch branching enzyme IIB; SBEIIB; corn; maize; antisense;

KW amylopectin; transgenic plant; pBE240; ss.

OS Zea mays.

FN Key Location/Qualifiers

FT CDS 79..2478

FT /\*tag- a

PN WO9722703-A2.

PD 26-JUN-1997.

PF 12-DEC-1996; 96WO-US19678.

PR 20-DEC-1995; 95US-0009113.

PA (DUPD) DU PONT DE NEMOURS & CO E I.

PI Broglie KE, Hubbard NL, Klein TM;

DR WPI: 1997-341694/31.

DR P-PADB; AAT69729.

PT Transgenic corn in which grain derived starch fine structure is

PT controlled - specifically branch chain distribution of amylopectin,

PT useful in preparation of thickened foodstuff

PS Example 1; Page 50-53; 92pp; English.

XX A CDNA insert (AAT69729) in plasmid clone pBE240 comprises a 2.7 kb  
 CC EcoRI-XhoI fragment isolated from a corn cDNA library. It includes  
 CC an open reading frame encoding starch branching enzyme IIB (SBEIIB)  
 CC (AAT69729). The insert was used as a starting point in the assembly  
 CC of DNA constructs (see also AAT69730, AAT69736-37) designed to  
 CC achieve suppression of SBEIIB expression in transgenic corn plants,  
 CC and thereby to produce novel starches that have properties  
 CC beneficial in food and industrial applications.

XX Sequence 2665 BP; 700 A; 525 C; 705 G; 735 T; 0 other;

Query Match 11.1%; Score 110.8; DB 18; Length 2665;

Best Local Similarity 88.0%; Pred. No. 8.2e-19;

Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 43 agatatgtatgatttcacgtcctgtagatag-gcttcaactcttgcatgacgtgcca 101  
 |||||  
 DB 1835 agatatgtatgatttcacgtcctgtagatagaccttcaactcttgcatgacgtgcca 1894  
 QY 102 tagcattacataaagatcaggtctgtcaccaatgggttaggtggaagctatctta 161  
 |||||  
 DB 1895 tagcattacataaagatcaggtctgtcaccaatgggttaggtggaagctatctta 1954  
 QY 162 actcatgggaatgagtttggtgacatccgt 191  
 |||||  
 DB 1955 attcatgggaatgagtttggtgacatccgt 1984

#### RESULT 14

AAV29757

ID AAV29757 standard; DNA; 2725 BP.

AC AAV29757;

DT 11-SEP-1998 (first entry)

DE Zea mays starch branching enzyme II gene.

KW SER; starch-encapsulating region; fusion vector;

KM starch branching enzyme II; ss.

OS Zea mays.

EH Key Location/Qualifiers

FT CDS 91..2490

FT /tag= a

FT sig\_peptide 91..264 /product= starch branching enzyme II

FT mat\_peptide 265..2487

FT /tag= b

FT /tag= c

PN MO9814601-A1.

PD 09-APR-1998.

PF 30-SEP-1997; 97WO-US17555.

PR 30-SEP-1996; 96US-0026655.

PA (EXSE-) EXSEED GENETICS LLC.

PI Guan H, Keeling P;

DR WPI: 1998-240100/21.

DR P-PSDB; AAW56489.

XX Hybrid polypeptide comprising starch-encapsulating region and

XX protein - useful for, e.g. producing protein(s) resistant to

XX degradation by stomach acids

XX Example 2; Page 44; 156pp; English.

XX The sequence is that of the starch branching enzyme II gene.

XX It can be used in the production of a hybrid polypeptide

XX comprising a starch-encapsulating region (SER) fused

XX to a payload protein. The hybrid polypeptide can be used to make

XX modified starches comprising the payload protein, selected from,

XX e.g. hormones, growth factors, antibodies, enzymes, dyes,

XX immunoglobulins, etc. The modified starch can also be used

CC to provide grain feeds enriched in amino acids. By encapsulating

CC the payload protein in starch, it is more resistant to

CC degradation by stomach acids.

XX Sequence 2725 BP; 727 A; 534 C; 715 G; 749 T; 0 other;

Query Match 11.1%; Score 110.8; DB 19; Length 2725;

Best Local Similarity 88.0%; Pred. No. 8.3e-19;

Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 43 agatatgtatgatttcacgtcctgtagatag-gcttcaactcttgcatgacgtgcca 101  
 |||||  
 DB 1847 agatatgtatgatttcacgtcctgtagatagaccttcaactcttgcatgacgtgcca 1906  
 QY 102 tagcattacataaagatcaggtctgtcaccaatgggttaggtggaagctatctta 161  
 |||||  
 DB 1907 tagcattacataaagatcaggtctgtcaccaatgggttaggtggaagctatctta 1966  
 QY 162 actcatgggaatgagtttggtgacatccgt 191  
 |||||  
 DB 1967 attcatgggaatgagtttggtgacatccgt 1996

#### RESULT 15

AAO73750

ID AAO73750 standard; cDNA to mRNA; 2919 BP.

AC AAO73750;

DT 05-JUL-1995 (first entry)

DE Rice starch branching enzyme.

KW Starch branching enzyme; rice; starch content; ds.

OS Oryza sativa.

EH Key Location/Qualifiers

FT 5'UTR 1..127

FT transit\_peptide 128..322

FT /tag= a

FT mat\_peptide 323..2606

FT /tag= b

FT /tag= c

FT /product= branching\_enzyme

FT 3'UTR 2603..2919

FT /tag= d

PN JP06261767-A.

PD 20-SEP-1994.

PF 22-OCT-1993; 93JP-0265171.

PR 29-OCT-1992; 92JP-0291719.

PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.

DR WPI: 1994-337418/42.

DR P-PSDB; AAR60811.

XX New gene of branching enzyme of rice starch - useful for

XX increasing starch yield of rice grains

XX Claim 1; Page 9-12; 13pp; Japanese.

XX The rice starch branching enzyme is encoded by the cDNA sequence

XX AAO73750. The starch content of rice grains can be increased by

XX increasing the expression of branching enzyme in rice plants.

XX Sequence 2919 BP; 755 A; 590 C; 800 G; 774 T; 0 other;

Query Match 10.6%; Score 106; DB 15; Length 2919;  
Best Local Similarity 86.0%; Pred. No. 1.5e-17;  
Matches 129; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 43 aggatatgtatgattcattcagctcttgatag-gcttcaactcttcgcatgtcgtygca 101  
|||||  
Db 1962 aggatatgtatgtatttatgctcttgacagacgcgcaacacctagcatgtatcgtygaa 2021  
|||||

QY 102 tagcattacataaataatgatcaggcttgcacatgggtttaggtgtgaaggctatctta 161  
|||||  
Db 2022 tagcattgcataaataatgatcaggcttgcacatgggtttaggtgtgaaggctatctta 2081  
|||||

QY 162 acttcatgggaaatgagttgggcatcctg 191  
|||||  
Db 2082 acttcatgggaaatgagttgggcatcctg 2111  
|||||

Search completed: September 12, 2002, 08:22:35  
Job time: 12312 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 08:12:56 ; Search time 92.79 Seconds  
(without alignments)  
2649.845 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_9500\_10500  
Perfect score: 1001  
Sequence: 1 tcgcagacatagacagcatga.....ctggagacatagctttttt 1001

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database :  
1: Issued\_Patents\_NA:\*  
2: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/lna/5B\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/lna/6A\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/lna/6B\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/lna/6C\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID                | Description       |
|------------|-------|-------------|--------|-------------------|-------------------|
| 1          | 110.8 | 11.1        | 2725   | US-08-941-445A-14 | Sequence 14, Appl |
| 2          | 94.2  | 9.4         | 3074   | US-09-087-277-1   | Sequence 1, Appl  |
| 3          | 92.6  | 9.3         | 1393   | US-09-087-277-3   | Sequence 3, Appl  |
| 4          | 60.4  | 6.0         | 3128   | US-08-716-449-1   | Sequence 1, Appl  |
| 5          | 59.2  | 5.9         | 2763   | US-08-941-445A-16 | Sequence 16, Appl |
| 6          | 46.2  | 4.6         | 2909   | US-08-104-158-1   | Sequence 1, Appl  |
| 7          | 40.6  | 4.1         | 2132   | US-08-844-188-39  | Sequence 39, Appl |
| 8          | 36.6  | 3.7         | 11478  | US-08-981-803-29  | Sequence 29, Appl |
| 9          | 36.6  | 3.7         | 11478  | US-08-983-440-29  | Sequence 29, Appl |
| 10         | 36.6  | 3.7         | 11478  | US-08-983-440-29  | Sequence 29, Appl |
| 11         | 36.4  | 3.6         | 2088   | US-09-351-414-3   | Sequence 3, Appl  |
| 12         | 36.4  | 3.6         | 2088   | US-09-351-414-3   | Sequence 3, Appl  |
| 13         | 35.8  | 3.6         | 5801   | US-08-647-351B-1  | Sequence 1, Appl  |
| 14         | 35.2  | 3.5         | 1945   | US-08-525-742-5   | Sequence 5, Appl  |
| 15         | 34.2  | 3.4         | 3418   | US-09-193-562D-29 | Sequence 29, Appl |
| 16         | 34.2  | 3.4         | 3418   | US-09-193-562D-29 | Sequence 29, Appl |
| 17         | 33.8  | 3.4         | 782    | US-08-261-825-1   | Sequence 1, Appl  |
| 18         | 33.8  | 3.4         | 782    | US-08-261-825-1   | Sequence 1, Appl  |
| 19         | 33.8  | 3.4         | 782    | US-08-261-825-1   | Sequence 1, Appl  |
| 20         | 33.8  | 3.4         | 782    | US-08-261-825-1   | Sequence 1, Appl  |
| 21         | 33.8  | 3.4         | 782    | US-08-261-825-1   | Sequence 1, Appl  |
| 22         | 33.8  | 3.4         | 782    | US-08-261-825-1   | Sequence 1, Appl  |
| 23         | 33.2  | 3.3         | 2368   | US-08-714-918-2   | Sequence 2, Appl  |
| 24         | 33.2  | 3.3         | 2368   | US-08-714-918-2   | Sequence 2, Appl  |
| 25         | 33.2  | 3.3         | 2368   | US-08-714-918-2   | Sequence 2, Appl  |
| 26         | 33.2  | 3.3         | 2368   | US-08-714-918-2   | Sequence 2, Appl  |
| 27         | 33.2  | 3.3         | 2368   | US-08-714-918-2   | Sequence 2, Appl  |

## ALIGNMENTS

```

RESULT 1
US-08-941-445A-14
; Sequence 14, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2725 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: mRNA
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 91..264
; FEATURE:
; NAME/KEY: mat_peptide

```

LOCATION: 265..2487  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 91..2490  
US-08-941-445A-14

Query Match 11.1%; Score 110.8; DB 3; Length 2725;  
Best Local Similarity 88.0%; Pred. No. 6.8e-22;  
Matches 132; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 43 aggatgatgatgattcagctctgtagtag-gctcaactcttcgcatgtagcgca 101  
DB 1847 AGGATGATGATGATTTCAAGCCCTCGATAGACCTTCACTACCATGATGATGCGGA 1906  
QY 102 tagcatataaataatgacagctgtccacatgggttagtggtgaagctatccta 161  
DB 1907 TAGCATATACATAAGATGATTAAGCTTATACCAATGGGTTTAGAGAGAGGCTATCTTA 1966  
QY 162 actcatgggaaatgagttgggcatcctg 191  
DB 1967 ATTTCATGGAAATGAGTTTGGACATCTTG 1996

RESULT 2  
US-09-087-277-1  
; Sequence 1, Application US/09087277B  
; Patent No. 6169226  
; GENERAL INFORMATION:  
; APPLICANT: EK, Bo  
; APPLICANT: KHOSNODI, Jamshid  
; APPLICANT: LARSSON, Clas-Tomas  
; APPLICANT: LARSSON, Hakan  
; APPLICANT: RASK, Lars  
; TITLE OF INVENTION: STARCH BRANCHING ENZYME II OF POTATO  
; FILE REFERENCE: 003300-486  
; CURRENT APPLICATION NUMBER: US/09/087,277B  
; EARLIER FILING DATE: 1998-05-29  
; EARLIER APPLICATION NUMBER: PCT/SE96/01558  
; EARLIER FILING DATE: 1996-11-28  
; EARLIER APPLICATION NUMBER: SE 9504272-7  
; EARLIER FILING DATE: 1995-11-29  
; EARLIER APPLICATION NUMBER: SE 9601506-0  
; EARLIER FILING DATE: 1996-04-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3074  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:beii gene  
; OTHER INFORMATION: (branching enzyme II) from Solanum tuberosum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (189)..(2825)  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (189)..(332)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (333)..(2825)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (92)..(2156)  
; OTHER INFORMATION: Nucleotides 92, 285, 1406, 1430, 1897 and 2156 are  
; OTHER INFORMATION: n wherein n = A, C, G or T.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (285)..(287)  
; OTHER INFORMATION: Amino acid -16 is Xaa wherein Xaa = Ile, Leu, Val  
; OTHER INFORMATION: or Phe.

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1404)..(1406)  
OTHER INFORMATION: Amino acid 358 is Xaa wherein Xaa = Leu or Phe.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1428)..(1430)  
OTHER INFORMATION: Amino acid 366 is Xaa wherein Xaa = Thr.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1896)..(1898)  
OTHER INFORMATION: Amino acid 522 is Xaa wherein Xaa = Tyr, Ser, Cys  
OTHER INFORMATION: or Phe.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (2154)..(2156)  
OTHER INFORMATION: Amino acid 608 is Xaa wherein Xaa = Pro.  
US-09-087-277-1

Query Match 9.4%; Score 94.2; DB 4; Length 3074;  
Best Local Similarity 78.1%; Pred. No. 3.5e-17;  
Matches 125; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 33 cctcttgcaagatatgatattcattcagctctgtagtag-gcttcaactcttcgcat 91  
DB 2112 cctgatggagaagatatgatatttattgctctgtagtagccttcaacatcataata 2171  
QY 92 gatctgtgcatgcatataataatgatcagctctgtacatggtttagtggtgaa 151  
DB 2172 gatcgtggtgcatgcatgacaaagatgtagctctgttaactatgtagtaggagaa 2231  
QY 152 ggcatacttaactatgggaaatgagttgggcatcctg 191  
DB 2232 gggtaoctaattcattcagtggaatgagatcgcgcaccccg 2271

RESULT 3  
US-09-087-277-3  
; Sequence 3, Application US/09087277B  
; Patent No. 6169226  
; GENERAL INFORMATION:  
; APPLICANT: EK, Bo  
; APPLICANT: KHOSNODI, Jamshid  
; APPLICANT: LARSSON, Clas-Tomas  
; APPLICANT: LARSSON, Hakan  
; APPLICANT: RASK, Lars  
; TITLE OF INVENTION: STARCH BRANCHING ENZYME II OF POTATO  
; FILE REFERENCE: 003300-486  
; CURRENT APPLICATION NUMBER: US/09/087,277B  
; EARLIER FILING DATE: 1998-05-29  
; EARLIER APPLICATION NUMBER: PCT/SE96/01558  
; EARLIER FILING DATE: 1996-11-28  
; EARLIER APPLICATION NUMBER: SE 9504272-7  
; EARLIER FILING DATE: 1995-11-29  
; EARLIER APPLICATION NUMBER: SE 9601506-0  
; EARLIER FILING DATE: 1996-04-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1393  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism:beii gene fragment  
; OTHER INFORMATION: (branching enzyme II) from Solanum tuberosum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)..(1393)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (424)..(1150)



```
;
; OTHER INFORMATION: Nucleotides 424, 891 and 1150 are n wherein n = A,
; OTHER INFORMATION: C, G or T.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (422)..(424)
; OTHER INFORMATION: Amino acid 141 is Xaa wherein Xaa = Thr.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (890)..(892)
; OTHER INFORMATION: Amino acid 297 is Xaa wherein Xaa = Tyr, Ser, Cys
; OTHER INFORMATION: or Phe.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1148)..(1150)
; OTHER INFORMATION: Amino acid 383 is Xaa wherein Xaa = Pro.
; US-09-087-277-3

Query Match          9.3%; Score 92.6; DB 4; Length 1393;
Best Local Similarity 77.5%; Pred. No. 7, 1e-17;
Matches 124; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 33 cctcttgcaaggatagtatgattcctgctcgtgataag-gcttcaactctgcgact 91
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1106 ctgattgacaagatagtatgatttcttgctctgctgataacnctcaatcattata 1165

QY 92 gatcgtgcatagcatataacatgatacagctgtgcacatgggttaagttgtgaa 151
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1166 gatcgtggtgatagcatgacgaagatgattgctgttaactatgattgaggaagaa 1225

QY 152 ggcatacttaactcattgggaatgagttgggcaccccg 191
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1226 gggtaacctaaattcattggaagaatgattcgccaccctg 1265

RESULT 4
US-08-716-449-1
; Sequence 1, Application US/08716449
; Patent No. 6103893
; GENERAL INFORMATION:
; APPLICANT: National Starch and Chemical Investment
; APPLICANT: Holding Corporation
; TITLE OF INVENTION: Method for Producing Altered Starch
; TITLE OF INVENTION: from Potato Plants
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott & Aylen
; STREET: Box 194, 24th Floor, Toronto-Dominion Bank Tower
; CITY: Toronto
; STATE: Canada M5K 1H6
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1MB 1.44 MB High Density Diskette
; COMPUTER: COMPAQ Deskpro 590 (IBM PC compatible)
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: Word 7.0 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,449
; FILING DATE: FILED CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00634
; FILING DATE: 22.03.95
; ATTORNEY/AGENT INFORMATION:
; NAME: Anita E. Nador
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1552
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 368-2400
; TELEFAX: (416) 363-7246
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3128 base pairs
```

```
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA encoding starch branching enzyme
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE: clone 1.2.1 and E2
; ORGANISM: Solanum tuberosum
; STRAIN: cv desiree
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE: mature tuber
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLER:
; IMMEDIATE SOURCE:
; LIBRARY: lambda Zap tuber cDNA
; CLONE: 1.2.1 and E2
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE: open reading frame
; NAME/KEY: starch branching enzyme
; LOCATION: 44-2788
; IDENTIFICATION METHOD: lone ORF with homology to other starch
; IDENTIFICATION METHOD: branching enzymes
; OTHER INFORMATION: complements KV832 E. coli glycogen
; OTHER INFORMATION: branching enzyme mutant
; US-08-716-449-1

Query Match          6.0%; Score 60.4; DB 3; Length 3128;
Best Local Similarity 69.5%; Pred. No. 1, 2e-07;
Matches 82; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 74 gctcaactcttcgcatgctggtgcatagcatatcataaattgacgctgtcacc 133
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1781 GAGCCTTCCTCCCTGTGTGATGAGGAAATGGCGCTTCACAGAATGATTCATTTCACA 1840

QY 134 atgggttttagtggtgaaggctatcttaacttcattggaatgagtttggcactcg 191
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1841 ATGGCCTTGAGAGAGAGGAGTACCTCAATTCATGCTGTAACGAGTTGCCATTCCTG 1898
```

```
RESULT 5
US-08-941-445A-16
; Sequence 16, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanning
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
```

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  US 60/026,855
3      FILING DATE:  30-SEP-1996
4      ATTORNEY/AGENT INFORMATION:
5      NAME:  WINNER, Ellen P
6      REGISTRATION NUMBER:  28,547
7      REFERENCE/DOCKET NUMBER:  89-97
8      TELECOMMUNICATION INFORMATION:
9      TELEPHONE:  (303) 499-8080
10     TELEFAX:  (303) 499-8089
11     INFORMATION FOR SEQ ID NO:  16:
12     SEQUENCE CHARACTERISTICS:
13     LENGTH:  2763 base pairs
14     TYPE:  nucleic acid
15     STRANDEDNESS:  single
16     TOPOLOGY:  not relevant
17     MOLECULE TYPE:  mRNA
18     HYPOTHETICAL:  NO
19     ORIGINAL SOURCE:
20     ORGANISM:  zea mays
21     FEATURE:
22     NAME/KEY:  transit_peptide
23     LOCATION:  2..190
24     FEATURE:
25     NAME/KEY:  mat_peptide
26     LOCATION:  191..2467
27     FEATURE:
28     NAME/KEY:  CDS
29     LOCATION:  2..2470
30     OS-08-941-445A-16

```

|                       |              |                  |               |             |
|-----------------------|--------------|------------------|---------------|-------------|
| Query Match           | 5.9%         | Score 59.2       | DB 3          | Length 2763 |
| Best Local Similarity | 68.3%        | Pred.No. 2.6e-07 |               |             |
| Matches 82            | Conservative | 0                | Mismatches 38 | Indels 0    |
|                       |              |                  |               | Gaps 0      |

[illegible]

RESULT 6  
US-08-104-158-1  
Sequence 1, Application US/08104158  
Patent No. 6215042  
GENERAL INFORMATION:  
APPLICANT: Willmitzer, Lothar  
APPLICANT: Sonnemwald, Uwe  
APPLICANT: Kossmann, Jens  
APPLICANT: Mueller-Roeber, Bernd  
APPLICANT: Visser, Richard Gerardus Franciscus  
APPLICANT: Jacobsen, Evert  
TITLE OF INVENTION: PLASMIDS CONTAINING DNA-SEQUENCES THAT  
TITLE OF INVENTION: CAUSE CHANGES IN THE CARBOHYDRATE CONCENTRATION AND THE  
TITLE OF INVENTION: CARBOHYDRATE COMPOSITION IN PLANTS, AS WELL AS PLANT CELLS  
TITLE OF INVENTION: AND PLANTS CONTAINING THESE PLASMIDS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen  
STREET: 1180 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-8403  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

1 CURRENT APPLICATION DATA:  
2 APPLICATION NUMBER: US/08/104,158  
3 FILING DATE: 13-AUG-1993  
4 CLASSIFICATION: 800  
5 PRIOR APPLICATION DATA:  
6 APPLICATION NUMBER: EP PCT/EP92/00302  
7 FILING DATE: 11-FEB-1992  
8 PRIOR APPLICATION DATA:  
9 APPLICATION NUMBER: DE P 41 04 782.6  
10 FILING DATE: 13-FEB-1991  
11 ATTORNEY/AGENT INFORMATION:  
12 NAME: Mellman, Edward A.  
13 REGISTRATION NUMBER: 24,735  
14 REFERENCE/DOCKET NUMBER: EA-1996 PCT (951-911  
15 TELECOMMUNICATION INFORMATION:  
16 TELEPHONE: 212-362-0700  
17 TELEFAX: 212-382-0888  
18 TELEX: 236925

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2909 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Solanum tuberosum cv. Desiree  
 STRAIN: Desiree  
 DEVELOPMENTAL STAGE: growing tuber  
 TISSUE TYPE: tuber  
 CELL TYPE: total tuber  
 IMMEDIATE SOURCE:  
 LIBRARY: cDNA of total tuber mRNA in pUC 19 (Hinc II)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2..1699  
 OTHER INFORMATION:  
 OTHER INFORMATION: (partial) truncated protein, 97,11 % identity to  
 OTHER INFORMATION: active potato branching enzyme"  
 IS-08-104-158-1

```

Query March 4.6%: Score 46.2: DB 4: Length 2909;
Best Local Similarity 67.2%: Pred. No. 0.0012:
Matches 80: Conservative 0: Mismatches 36: Indels 1: Gaps 1:

QY 74 gcttaacctctgcgattgatcgtgcatagattacataaattgat-caggtctgtcac 132
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1601 GATGTTGTCCTCGTGTGTGATGCACAGAAATTGGCGCTTGACAAAGATTCATTTTTTCAC 1660

QY 133 catgggttttaggttggtgaaggtcatcttaactcatgggaattgaagtttggcatcctg 191
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1661 AATGGCCTTGGGAGAGAGAGGGGTACTCATTTTCATGGGTGATACAGATTGGCCATCCTG 1719

RESULT 7
US-08-844-188-39/C
; Sequence 39, Application US/08844188
; Patent No. 6127180
; GENERAL INFORMATION:
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Pesticidal Toxins
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1

```

```

; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/844.188
; APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/633,993
; FILING DATE: 19-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-703C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2132 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-844-188-39

```

```

Query Match      4.1%; Score 40.6; DB 3; Length 2132;
Best Local Similarity 51.4%; Pred. No. 0.041;
Matches 94; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 490 atgaatcataatgctttagaagaatcaacatgctttagttagcgtta 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2009 ATACGATTAATCTTCAAAATTAATTAAGATTAATTAATTAATTTGTTGTTAGATTAAT 1950
QY 550 acaataagatgcttgagagctgctgacatcaaaatacatgatttttgcaggaga 609
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1949 TCTAATTTCAATTTTAAAAAATATATTTTAAATTTTAAAGTACTTTTAAAGATTA 1890
QY 610 tgcagattcttagatcatcgtagatgcaagagctgacgcagcaatgcacatctga 669
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1889 TTTGTTATTTCTTCACTCTTCATATGATTTGTAAGAATTAATTAAGCTTGTGA 1830
QY 670 gga 672
    ||
DB 1829 TGA 1827

```

```

RESULT 8
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEFFELINGER, F.
; APPLICANT: PARKER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/232.463
; APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US/07/935,313
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

```

```

Query Match      3.7%; Score 36.6; DB 1; Length 7218;
Best Local Similarity 5.4%; Pred. No. 0.95;
Matches 18; Conservative 174; Mismatches 143; Indels 0; Gaps 0;

QY 154 ctatcttaactcatcggaatagtggtgcatctgtagcttcttacacattatg 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1034 CGAGCTTGCGCTGACGCTGAGGAGCTTGCGATTTTTTTTTTTTTTTTTTTTTT 1093
QY 214 catctgatgatgctgtagtcttactglaattgaaccatgcttctctacatgtagt 273
    : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1094 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1153
QY 274 attatgatactgctgcttcccaaggagaagtaactcttacttctctgccaatgtagt 333
    : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1154 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1213
QY 334 agatttccaaggagcccaaacactctccaacgcgcaagctctccctggaataca 393
    : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1214 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1273
QY 394 atagtaataaalcgcgcgtagattgactgtagttagctgtagctgtagctatgtagt 453
    : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1274 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1333
QY 454 tccctcaactagatcttattgagcattattctt 488
    : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1334 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1368

```

```

RESULT 9
US-08-981-803-29
; Sequence 29, Application US/08981803
; Patent No. 6147279
; GENERAL INFORMATION:
; APPLICANT: FOULSEN, PETER
; TITLE OF INVENTION: INHIBITION OF GENE EXPRESSION
; FILE REFERENCE: 674509-2004
; CURRENT APPLICATION NUMBER: US/08/981,803
; CURRENT FILING DATE: 1997-04-17
; EARLIER APPLICATION NUMBER: PCT/EP96/03052
; EARLIER FILING DATE: 1996-07-12

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

|                          |        |                |           |               |
|--------------------------|--------|----------------|-----------|---------------|
| Query Match              | 3.7%;  | Score 36.6;    | DB 3;     | Length 11478; |
| Best Local Similarity    | 65.1%; | Pred. No. 1.2; |           |               |
| Matches 54; Conservative | 0;     | Mismatches 29; | Indels 0; | Gaps 0;       |

| Qy   | Db  | Qy                        | Db                        |
|--|---|---------------------------|---------------------------|
| 107  | 6667  | 167                       | 6727                      |
| ttacaataaattgcatcagccttgcaccacttgggtttagtgcgtgaagcgcattctaac | taaccacagatgatccatttttccaacaagcgctcttggagagaggttactcctaattc | atgggaatatgagtttgggcattcc | atgggttaacgagtgatgtcttacc |
| 166  | 6726  | 189                       | 6749                      |
|  |   |                           |                           |

```

RESULT 10
US-08-983-440-29
Sequence 29, Application US/08983440
Patent No. 6232122
GENERAL INFORMATION:
APPLICANT: POLISEN, Peter
TITLE OF INVENTION: INHIBITION OF GENE EXPRESSION
FILE REFERENCE: 674509-2003
CURRENT APPLICATION NUMBER: US/08/983,440
CURRENT FILING DATE: 1998-04-17
EARLIER APPLICATION NUMBER: 9514437.4
EARLIER FILING DATE: 1995-07-14
EARLIER APPLICATION NUMBER: PCT/EP96/03053
EARLIER FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 29
LENGTH: 11478
TYPE: DNA
ORGANISM: Solanum tuberosum
FEATURE:
NAME/KEY: variation
LOCATION: (1)..(11478)
OTHER INFORMATION: B stands for G or C or T/U
FEATURE:
NAME/KEY: variation
LOCATION: (1)..(11478)
OTHER INFORMATION: R stands for G or A
FEATURE:
NAME/KEY: variation
LOCATION: (1)..(11478)
OTHER INFORMATION: K stands for G or T/U
FEATURE:
NAME/KEY: variation
LOCATION: (1)..(11478)
OTHER INFORMATION: W stands for A or T/U
FEATURE:
NAME/KEY: variation
LOCATION: (1)..(11478)
OTHER INFORMATION: M stands for A or C
US-08-983-440-29

```

|                       |  |                 |                |               |
|-----------------------|--|-----------------|----------------|---------------|
| Query Match           | 3.7%   | Score 36.6;     | DB 4;          | Length 11478; |
| Best Local Similarity | 65.1%  | Pred. No. 1.2;  |                |               |
| Matches               | 54;  | Conservative 0; | Mismatches 29; | Indels 0;     |
| Gaps                  | 0;   |                 |                |               |
| Q7_107                | ttactataatgatacaggctctgcaccatgggttagtggtggaagcgcatcttaattc | 166             |                |               |
|                       |  |                 |                |               |

Db 6667 taaccacagatgatccatttttttccaacaatgaccttgagagagaggggttaactcaatttc 6726

Qy 167 atgggaatgatgttttgcaatcc 169  
||||| || ||| | | | |

Db 6727 atgggtaacgaggtatgtctcttac 6749

```

RESULT 11
US-09-351-414-3
Sequence 3, Application US/09351414
Patent No. 6265199
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Baldwin, Nard
APPLICANT: Baisner, Theresa A.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: DISINTEGRIN HOMOLOG
FILE REFERENCE: 98-29
CURRENT APPLICATION NUMBER: US/09/351,414
CURRENT FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FASTSEQ for Windows Version 3.0.
SEQ ID NO 3

```

```

? ORGANISM: Artificial Sequence
?
? FEATURE:
? OTHER INFORMATION: zdm1l amino acid degenerate sequence
?
? NAME/KEY: variation
? LOCATION: (1)..(2088)
? OTHER INFORMATION: n is any nucleotide
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)..(2088)
? OTHER INFORMATION: n = A,T,C or G
US-09-351-414-3

```

|    |                       |   |                 |        |                                   |
|----|-----------------------|---|-----------------|--------|-----------------------------------|
|    | Query Match           | 3.6%  | Score 36.4      | DB 4   | Length 2086                       |
|    | Best Local Similarity | 29.3%   | Pred. NO. 0.63, |        |                                   |
|    | Matches               | 60  | Conservative    | 38     | Mismatches 107; Indels 0; Gaps 0; |
| Oy | 237                   | tgtatttgaacacagcttcctcctccacatgtaatagtatgaatactgtgtgccaa    | 296             |        |                                   |
|    |                       | :     :   | :               | :    : | :                                 |
| Dd | 1006                  | tgyethagsgaragaracngnngltnsncaeywmgnaartlywnsaartgyysnathyn | 1065            |        |                                   |
| Oy | 237                   | ggagaagttaactctattactttgcagaatgatataattcccagaaggccacaaa     | 356             |        |                                   |
|    |                       | : :   :   :   :   | :               | :   :  | :                                 |
| Dd | 1066                  | garlaymngayltvyltnrcarnngnngnngncntgyrntlyaa yngnccnaear    | 1125            |        |                                   |
| Oy | 357                   | cctctccaacgcgcgaagtctccctcccggaataacaatagtagtaataatgcgcgcgt | 416             |        |                                   |
|    |                       | :     :     :   | :               | :      | :                                 |
| Dd | 1126                  | ynttygarccnacnagrttgyynaagyntatyngarngcnngargartgyyalygy    | 1185            |        |                                   |
| Oy | 417                   | agatttgactctgttaagtttagct                                   | 441             |        |                                   |
|    |                       | :   :   :   :   |                 |        |                                   |
| Dd | 1186                  | ggnntyacyglngartgytaygny                                    | 1210            |        |                                   |

RESULT 12  
US-08-647-351B-1  
Sequence 1, Application US/08647351B  
Patent No. 5770368  
GENERAL INFORMATION:  
APPLICANT: De Leon, Ricardo  
APPLICANT: Rochelle, Paul  
TITLE OF INVENTION: Cryptosporidium Detection Method  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 S. Lake Avenue, 9th Floor  
CITY: Pasadena

```

; STATE: California
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,351B
; FILING DATE: May 9, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Farah, David A.
; REGISTRATION NUMBER: 38,134
; REFERENCE/DOCKET NUMBER: 11364
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 796-4000
; TELEFAX: (626) 795-6321
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-08-647-351B-1

```

```

Query Match      3.6%; Score 36; DB 1; Length 3607;
Best Local Similarity 55.6%; Pred. No. 1;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```

```

QY 477 cattattcttgatgaatcattatgtttaggaagaatcattgctttagt 536
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 91 CATTGATTTTCATTAATAATATCATATGATTGATTGCTGTAATAATATTGATTGAT 150
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 537 ttctgaagcgttaacataagatctgttgagagcttgatcattaaataatcattgatt 596
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 151 GAGGAGATCATATTCACCTTTATTATTAGAAATTTTATTAAATAAAGTTGATTATT 210
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 597 tttt 600
      |||||
DB 211 TTTT 214

```

```

RESULT 13
US-09-177-249-5/c
; Sequence 5, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Onad, NLR
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Yadeagari, Ramin
; APPLICANT: Margossian, Linda
; APPLICANT: Harada, John
; APPLICANT: Goldberg, Robert B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
; FILE REFERENCE: 023070-08612005
; CURRENT APPLICATION NUMBER: US/09/177,249
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: US 09/071,838
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 5801
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS

```

```

; LOCATION: (3872)..(5566)
; OTHER INFORMATION: fertilization-independent endosperm 3 (FIE3)
; OTHER INFORMATION: WD40/polycomb gene genomic sequence
; US-09-177-249-5

```

```

Query Match      3.6%; Score 35.8; DB 4; Length 5801;
Best Local Similarity 50.9%; Pred. No. 1.5;
Matches 85; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```

```

QY 469 ttatggccattattcttgatgaatcattatgtttaggaagaatcattgct 528
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 256 TTAATTCATGCTGGTGTCTTAAGAATTTTACCGTTTGGAAATTTAATATGCTCT 197
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 529 ttctgaagcgttaacataagatctgttgagagcttgatcattaaata 588
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 196 TTTTCAAAATTAAGATCGATTTGTCATGACGAATCATGATTTGACGAATACACATTC 137
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 589 tcattgatttttcagaggagatcagatttcttgatcattgctgta 635
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 136 TCATGCTTATTGTTGCTCTCTTAATACGCTTACATATATATGTA 90
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 14
US-08-525-742-5
; Sequence 5, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Okawa, Setuko
; APPLICANT: Sasaki, Saki
; APPLICANT: Ohnawa, Ikuroh
; APPLICANT: Funato, Hiroko
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemasa
; APPLICANT: Takahashi, Kiyohiko
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLealand, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930

```

```

; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1945 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..1038
; US-08-525-742-5

```

```

Query Match          3.5%; Score 35.2; DB 2; Length 1945;
Best Local Similarity 47.7%; Pred. No. 1.3;
Matches 103; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

```

```

QY 384 ggaataacaataagatgatgataaagccgcgtagatttgatctgtgaagtttagctgt 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1433 GGATTTAAACGGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492

QY 444 gctattacatccctcactagatcttatttggccatttattcttgatgaaatcaatg 503
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1493 TTGAAGAGTGTTCATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1552

QY 504 ttittgagaaagatcaacatcttcttgtagttttgtagagcgttaacataagatggt 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1553 GTTATCTGAATAGATTTAGAGTCTTATTTAGTTGATCAACAATACCATCGTTCCTTT 1612

QY 564 tgagagtggttgatcatataaataatcatgatttt 599
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1613 CATTAAGCTCAGTAAGTAATGATTTTCAATCTT 1648

```

```

RESULT 15
US-09-193-562D-29
; Sequence 29, Application US/09193562D
; Patent No. 6309857
; GENERAL INFORMATION:
; APPLICANT: Pauli, Benedicht U.
; TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
; FILE OF INVENTION: Activated Chloride Channel-Adhesion Molecules
; CURRENT APPLICATION NUMBER: US/09/193,562D
; CURRENT FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US/60/065,922
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 29
; LENGTH: 3418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-193-562D-29

```

```

Query Match          3.4%; Score 34.2; DB 4; Length 3418;
Best Local Similarity 47.7%; Pred. No. 3.3;
Matches 133; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

```

```

QY 340 tcgaagagggccacaactctcccaaccgcaaggtctcccttgaataatacaatagtt 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2630 tcctagatgccatgctccctcagaggttcaaacatgcaagaaccaactattatctc 2689

QY 400 atgataatggcgcgtagatttgatctgtgaagtttagctgtgtattatccctc 459
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2690 ctccacaggaaccagatctccgtgctgggtaccacaattctgcacatcagttggcaa 2749

QY 460 actagatcttattggccattattcttggatgaatcatatgtttttaggaagatc 519
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2750 ttittgattagcgtgaatttatctatatttatactagaataatataatgaactcaa 2809

QY 520 aacatgctttttagtattttagtagagcgttaacataa---glatggttgagagttgtga 576
    ..

```

```

Db 2810 ttcaatgtaatacactactggtgaacatttatttaaaatttaactaactattgt 2869
QY 577 tcattaaaataatcatgattttttgcaggagagatgcaga 615
    ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2870 ctattataagctcatatataataataaagtgaaagtlaca 2908

```

Search completed: September 12, 2002, 08:13:33  
Job time: 14870 sec

---





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 08:10:22 : Search time 3207.55 Seconds  
(without alignments)  
4212.078 Million cell updates/sec

Title: US-09-508-377-10\_COPY\_9500\_10500

Perfect score: 1001

Sequence: 1 tccgacatagaacacataga.....ctggagacatagctttttt 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estlro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_lun:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 199.4 | 19.9        | 465    | 12    | A0690523 nbxb00820  |
| 2          | 136.4 | 13.6        | 650    | 10    | BE427534 PSR7153 I  |
| 3          | 133.2 | 13.3        | 244    | 10    | BE591091 WHE0861_C  |
| 4          | 133.2 | 13.3        | 375    | 10    | BE419958 WMS019.E6  |
| 5          | 133.2 | 13.3        | 499    | 10    | BE517857 WHE01803.A |
| 6          | 133.2 | 13.3        | 592    | 10    | BE402623 CSB009H05  |
| 7          | 130   | 13.0        | 471    | 9     | AV914376 AV914376   |
| 8          | 130   | 13.0        | 801    | 9     | BE195628 HVSMEH008  |
| 9          | 130   | 13.0        | 825    | 10    | BF065047 HV_CEB002  |
| 10         | 118.8 | 11.9        | 448    | 10    | D48464 RICS14669A   |
| 11         | 117.4 | 11.7        | 743    | 10    | BE414598 SCU011.E0  |
| 12         | 114.2 | 11.4        | 494    | 10    | BE498533 WHE0971_G  |
| 13         | 114.2 | 11.4        | 688    | 9     | AV944141 AV944141   |
| 14         | 114.2 | 11.4        | 692    | 9     | AV921312 AV921312   |
| 15         | 114.2 | 11.4        | 728    | 10    | BI948391 HVSMEI000  |
| 16         | 113.2 | 11.3        | 506    | 10    | BG366889 HVSMEI000  |
| 17         | 110.6 | 11.0        | 832    | 9     | BE196055 HVSMEH009  |

|    |       |      |     |    |          |                     |
|----|-------|------|-----|----|----------|---------------------|
| 18 | 110.6 | 11.0 | 918 | 9  | BE196057 | BE196057 HVSMEH009  |
| 19 | 106   | 10.6 | 243 | 9  | AA751373 | AA751373 ISYS0039   |
| 20 | 98.4  | 9.8  | 651 | 10 | BG526727 | BG526727 63-11 Ste  |
| 21 | 96.4  | 9.6  | 279 | 10 | BI072226 | BI072226 C071P70U   |
| 22 | 96.4  | 9.6  | 372 | 10 | BI069710 | BI069710 C006P40U   |
| 23 | 95.6  | 9.6  | 565 | 9  | A1795537 | A1795537 605011A01  |
| 24 | 95.4  | 9.5  | 731 | 10 | BE602527 | BE602527 HVSMEH009  |
| 25 | 92.8  | 9.3  | 548 | 9  | AW258093 | AW258093 68706BD07  |
| 26 | 91.2  | 9.1  | 168 | 10 | BM259431 | BM259431 952009B08  |
| 27 | 91.2  | 9.1  | 560 | 9  | A1737162 | A1737162 605030C06  |
| 28 | 91.2  | 9.1  | 583 | 9  | A1737050 | A1737050 605017B02  |
| 29 | 91.2  | 9.1  | 592 | 9  | AW231887 | AW231887 687060C09  |
| 30 | 91.2  | 9.1  | 612 | 9  | AW061921 | AW061921 687009B05  |
| 31 | 91.2  | 9.1  | 613 | 9  | A1665623 | A1665623 605013C06  |
| 32 | 91.2  | 9.1  | 615 | 9  | A1759057 | A1759057 605084D05  |
| 33 | 90.2  | 9.0  | 526 | 9  | AW216035 | AW216035 687050F06  |
| 34 | 90    | 9.0  | 433 | 10 | BG551002 | BG551002 sad31h09.  |
| 35 | 90    | 9.0  | 582 | 10 | BM309136 | BM309136 sak55b02.  |
| 36 | 90    | 9.0  | 689 | 10 | BI308555 | BI308555 EST529965  |
| 37 | 89.6  | 9.0  | 595 | 9  | A1833354 | A1833354 605088C12  |
| 38 | 89    | 8.9  | 621 | 10 | BG521671 | BG521671 13-3 Stev  |
| 39 | 88.4  | 8.8  | 509 | 10 | BG651974 | BG651974 sad72d10.  |
| 40 | 88.4  | 8.8  | 571 | 10 | BI465105 | BI465105 sai107e09. |
| 41 | 88.4  | 8.8  | 571 | 10 | BI699680 | BI699680 sai125g09. |
| 42 | 88.4  | 8.8  | 589 | 9  | AW596011 | AW596011 s196g03.Y  |
| 43 | 86.8  | 8.7  | 601 | 9  | A1993785 | A1993785 701514536  |
| 44 | 85.4  | 8.5  | 184 | 9  | AA751420 | AA751420 ISYS0089   |
| 45 | 85.2  | 8.5  | 569 | 9  | AV442128 | AV442128 AV442128   |

## ALIGNMENTS

RESULT 1  
A0690523/c 465 bp DNA linear GSS 01-JUL-1999  
LOCUS nbxb0082011f, DNA sequence.

DEFINITION A0690523.1 GI:5331691  
ACCESSION A0690523  
VERSION A0690523.1  
KEYWORDS GSS.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa

REFERENCE Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 465)  
Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293

FEATURES  
Source Email: rwing@clemson.edu  
Seq primer: TATACGACTCATTATAGCG  
Class: BAC ends  
High quality sequence stop: 177.  
Location/Qualifiers  
1..465  
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nbxb0082011f"  
/clone.lib="CUGI Rice BAC Library"  
/tissue\_type="leaf"  
/lab\_host="E. coli DH10B"  
/note="Vector: pBel08C11; Site\_1: HindIII; Site\_2:  
HindIII; Rice is one of two most popular grains in the  
world. Half of the world population especially those



Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
Seq primer: Stragatene SK primer.

## FEATURES

source

Location/Qualifiers

1..244  
/organism="Triticum aestivum"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="WHE0861\_C09\_F17"  
/clone\_1lb="wheat 20-45 DAP spike cDNA library"  
/tissue\_type="Spike and seed"  
/dev\_stage="Adult plant"  
/lab\_host="E. coli SOLR"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; site:1. EcoRI; site:2. XhoI; Plants were grown in the greenhouse. Spikes at 20 DAP and seeds at 30 to 45 DAP were harvested, total RNA and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the T7 Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors )."

BASE COUNT 66 a 41 c 64 g 73 t  
ORIGIN

## Query Match

Best Local Similarity 97.3%; Score 133.2; DB 10; Length 244;  
Matches 146; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 43 aggatgtatgatttcacgtccttgatag-gctcaactcttcgatgcgtggca 101  
|||||  
Db 86 AGGATGTGATGATTCATGCTCTGTGATGACCTTCACCTCGCATGATCGGCA 145  
Qy 102 tagcattacataaagatcaggtctgtcaccatgggttaggttggaagctatcta 161  
|||||  
Db 146 TAGCATTACATTAAGATGATGAGCTGTGTCACCATGGGTTAGTGGCAAGCTATCTTA 205  
Qy 162 acttcacgggaatgagttgggcatcctg 191  
|||||  
Db 206 ACTTCATGGGAATGAGTTGGGCATCTCG 235

## RESULT 4

BEA19958

LOCUS BEA19958 375 bp mRNA linear EST 24-JUL-2000  
DEFINITION WMS019.E6R000101 ITTC WMS Wheat Scutellum library Triticum aestivum  
CDNA clone WMS019.E6, mRNA sequence.

ACCESSION BEA19958  
VERSION BEA19958.1 GI:9417804

KEYWORDS EST  
SOURCE bread wheat.

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticaceae; Triticum.

REFERENCE 1 (bases 1 to 375)

AUTHORS

Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier,  
S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,  
Herrmann,R.G., Holton,T., Jacquemais,J.M., Jia,J., Joudrier,P.,  
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogilshay,Y.,  
Pechioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shafitlou,M.,  
Sorrells,M., Warburton,M. and Wenzel,G.  
International Triticaceae EST Cooperative (ITTC): Production of  
Expressed Sequence Tags for Species of the Triticaceae  
Unpublished (2000)

JOURNAL COMMENT  
Contact: Schuch W

zeneca Wheat Improvement Centre, Norwich Research Park  
Colney Lane, Norwich NR4 7UH UNITED KINGDOM  
Tel: 44 1603 250 2600  
Fax: 44 1603 250 699  
Email: wolfgang.schuch@guk.zeneca.com

International Triticaceae EST Cooperative (ITTC)  
http://wheat.pw.usda.gov/genome.  
Location/Qualifiers

## FEATURES

source

Location/Qualifiers

1..375  
/organism="Triticum aestivum"  
/cultivar="Novosibirskaya 67"  
/db\_xref="taxon:4565"  
/clone="WMS019\_E6"  
/clone\_1lb="ITTC WMS Wheat Scutellum library"  
/tissue\_type="scutellum callus"  
/note="M13 Reverse sequencing primer used for 5' end of clone."

BASE COUNT 107 a 70 c 89 g 108 t 1 others  
ORIGIN

## Query Match

Best Local Similarity 97.3%; Score 133.2; DB 10; Length 375;  
Matches 146; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 43 aggatgtatgatttcacgtccttgatag-gctcaactcttcgatgcgtggca 101  
|||||  
Db 83 AGGATGTGATGATTCATGCTCTGTGATGACCTTCACCTCGCATGATCGGCA 142  
Qy 102 tagcattacataaagatcaggtctgtcaccatgggttaggttggaagctatcta 161  
|||||  
Db 143 TAGCATTACATTAAGATGATGAGCTGTGTCACCATGGGTTAGTGGCAAGCTATCTTA 202  
Qy 162 acttcacgggaatgagttgggcatcctg 191  
|||||  
Db 203 ACTTCATGGGAATGAGTTGGGCATCTCG 232

## RESULT 5

BE517857

LOCUS BE517857 499 bp mRNA linear EST 08-AUG-2000  
DEFINITION WHE0803\_A06\_B11ZS Wheat vernalized crown cDNA library Triticum  
aestivum cDNA clone WHE0803\_A06\_B11, mRNA sequence.

ACCESSION BE517857  
VERSION BE517857.1 GI:9741887

KEYWORDS EST  
SOURCE bread wheat.

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticaceae; Triticum.

REFERENCE 1 (bases 1 to 499)

AUTHORS

Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han  
P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,  
Seaton,C.L. and Tong,J.C.  
The structure and function of the expressed portion of the wheat  
genomes - Vernalized crown cDNA library  
Unpublished (2000)

JOURNAL COMMENT  
Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 51053585773  
Fax: 51053595818

REFERENCE 1 (bases 1 to 499)

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
Seq primer: Stragatene SK primer.  
Location/Qualifiers

## FEATURES

source

Location/Qualifiers

1..499  
/organism="Triticum aestivum"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="WHE0803\_A06\_B11"  
/clone\_1lb="Wheat vernalized crown cDNA library"  
/tissue\_type="Crown tissue of seedling"  
/dev\_stage="Five-week old seedling"  
/lab\_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site\_1: EcoRI; Site\_2: XhoI; Seeds were germinated and grown at 4 C for 5 weeks. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give Bluescript phagemids in the T7 Close Lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 149 a 91 c 121 g 138 t

Query Match 13.3%; Score 133.2; DB 10; Length 499;  
Best Local Similarity 97.3%; Pred. No. 1.3e-20;  
Matches 146; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Oy 43 aggatattgatattcattgctctgtagtag-gctcaactcttcgattgtagtgcga 101  
|||||  
Db 105 AGGATATGATGATTTCATGGCTCTGGATAGACCTCAACCTTCGCAATTCGTGGCA 164  
|||||  
Oy 102 tagcatcataaataatgatacagctgtcacatggtttgagtgtaagctactcta 161  
|||||  
Db 165 TAGCATTCATTAATAATGATCAGGCTTGACCATGGCTTTAGTGCGAAGGCTATCTTA 224  
|||||

Oy 162 actcatgggaatgattggtggacacctg 191  
|||||  
Db 225 ACTCATGGGAATGAGTTGGGCAATCCTG 254  
|||||

RESULT 6  
BE402623 592 bp mRNA linear EST 21-JUL-2000  
LOCUS CSB009H05990908 ITFC CSB Wheat Endosperm Library Triticum aestivum  
DEFINITION cDNA clone CSB009H05, mRNA sequence.  
ACCESSION BE402623  
VERSION BE402623.1 GI:9362091  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae  
1 (bases 1 to 592)

REFERENCE  
AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.-J., McGuire, P., Ogihara, T., Pecchioni, N., Qualset, C., Schuch, W., Selvaraj, G., Shariflou, M., Sorrells, M., Warburton, M., and Wenzel, G.  
International Triticaceae EST Cooperative (ITEC): Production of Expressed Sequence Tags for Species of the Triticaceae  
Unpublished (2000)

JOURNAL  
COMMENT Contact: Appels R  
Div. of Plant Industry, CSIRO  
Canberra ACT 2601 AUSTRALIA  
Tel: 61 62 465496  
Fax: 61 62 465000  
Email: rudiepi.csiro.au  
International Triticaceae EST Cooperative (ITEC)  
http://wheat.pw.usda.gov/genome.  
Location/Qualifiers

FEATURES  
source 1..592  
/organism="Triticum aestivum"  
/cultivar="Wynna"  
/db\_xref="taxon:4565"  
/clone\_id="ITFC CSB Wheat Endosperm Library"  
/tissue\_type="endosperm"  
/dev\_stage="8-12 days post anthesis"  
/lab\_host="Escherichia coli SOLR"  
/note="Vector: Lambda Zap/Bluscript; Site\_1: XhoI; Site\_2: EcoRI; Plants grown in Phytotron with 18C/13C

(day/night) 16 hour light. M13 Reverse sequencing primer used. 1.0 kbp average insert size."

BASE COUNT 164 a 116 c 144 g 167 t 1 others

Query Match 13.3%; Score 133.2; DB 10; Length 592;  
Best Local Similarity 97.3%; Pred. No. 1.3e-20;  
Matches 146; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Oy 43 aggatattgatattcattgctctgtagtag-gctcaactcttcgattgtagtgcga 101  
|||||  
Db 6 AGGATATGATGATTTCATGGCTCTGGATAGACCTCAACCTTCGCAATTCGTGGCA 65  
|||||  
Oy 102 tagcatcataaataatgatacagctgtcacatggtttgagtgtaagctactcta 161  
|||||  
Db 66 TAGCATTCATTAATAATGATCAGGCTTGACCATGGCTTTAGTGCGAAGGCTATCTTA 125  
|||||  
Oy 162 actcatgggaatgattggtggacacctg 191  
|||||  
Db 126 ACTCATGGGAATGAGTTGGGCAATCCTG 155  
|||||

RESULT 7  
AV914376 471 bp mRNA linear EST 18-JAN-2002  
LOCUS AV914376 K. Sato unpublished cDNA library, cv. Haruna Nijo  
DEFINITION germination shoots Hordeum vulgare subsp. vulgare cDNA clone  
bgs5011 5', mRNA sequence.  
ACCESSION AV914376  
VERSION AV914376.1 GI:18210153  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare.  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae  
1 (bases 1 to 471)

REFERENCE  
AUTHORS Sato, K., Saitoh, D. and Takeda, K.  
TITLE Barley EST sequencing project in NIG and Okayama Univ  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-1  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yatai, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp.  
Location/Qualifiers

FEATURES  
source 1..471  
/organism="Hordeum vulgare subsp. vulgare"  
/cultivar="Haruna Nijo"  
/db\_xref="taxon:112509"  
/clone="bgs5011"  
/clone\_id="K. Sato unpublished cDNA library, cv. Haruna Nijo germination shoots"  
/tissue\_type="shoots"  
/dev\_stage="germination"

BASE COUNT 138 a 81 c 116 g 135 t 1 others

Query Match 13.0%; Score 130; DB 9; Length 471;  
Best Local Similarity 96.0%; Pred. No. 6.9e-20;  
Matches 144; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Oy 43 aggatattgatattcattgctctgtagtag-gctcaactcttcgattgtagtgcga 101  
|||||  
Db 100 AGGATATGATGATTTCATGGCTCTGGATAGACCTCAACCCCTCGCATTCGTGGCA 159  
|||||  
Oy 102 tagcatcataaataatgatacagctgtcacatggtttgagtgtaagctactcta 161  
|||||  
Db 160 TAGCATTCATTAATAATGATCAGGCTTGACCATGGCTTTAGTGCGAAGGCTATCTTA 219  
|||||

QY 162 acttcattggaatgagttgggcatctg 191  
 Db 220 ATTTCATGGGAATGAGTTGGGCATCTTG 249

RESULT 8

LOCUS BE195628 801 bp mRNA linear EST 22-OCT-2001  
 DEFINITION HVSMNH0089P20f Hordeum vulgare 5-45 DAP spike EST library  
 HVCNMA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMNH0089P20f,  
 mRNA sequence.  
 BE195628  
 VERSION BE195628.2 GI:13188305  
 KEYWORDS EST.  
 SOURCE barley.  
 ORGANISM Hordeum vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
 ; Triticeae; Hordeum.  
 1 (bases 1 to 801)  
 Wing,R., Close,T.J., Kleinofs,A., Wise,R., Begum,D., Frisch,D., Yu  
 ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton  
 ,R.D., Close,S.J., Oates,R. and Main,D.  
 Development of a genetically and physically anchored EST resource  
 for barley genomics: Morex 5-45 DAP spike cDNA library  
 Unpublished (2001)  
 On Jun 26, 2000 this sequence version replaced gi:8707823.  
 CONTACT: Wing RA  
 CLEMSON UNIVERSITY GENOMICS INSTITUTE  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 TEL: 864 656 7288  
 FAX: 864 656 4293  
 EMAIL: rwing@clermson.edu  
 Total hg bases = 408  
 Seq primer: AATTAACTCTCACTAAAGCG  
 High quality sequence stop: 611.  
 Location/Qualifiers  
 1..801  
 /organism="Hordeum vulgare"  
 /cultivar="Morex"  
 /db\_xref="taxon:4513"  
 /clone\_lib="HVSMNH0089P20f"  
 HVCNMA0009 (5 to 45 DAP)"  
 /tissue\_type="5-45 DAP Spike"  
 /lab\_host="SOLR"  
 /note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI;  
 plants were grown in the greenhouse at the University of  
 California, Riverside (Fenton, SJ Close, TJ Close). Whole  
 spikes with awns trimmed were collected at 5, 10, 15, 20,  
 30 and 45 DAP (Fenton). Total RNA was prepared from each  
 pool, equal quantities of all six RNA pools were combined,  
 poly(A) RNA was purified from the mixture, one primary  
 unamplified cDNA library was made, and 1 million pfu were  
 in vivo excised to give Bluescript SK(-) cDNA phagemids  
 (Choi) in the TJ Close lab at the University of California,  
 Riverside. Phagemids were plated and picked at the Clemson  
 University Genomics Institute (CUGI) (Begum, Palmer,  
 Frisch, Atkins and Wing). Plasmid DNA preparations, DNA  
 sequencing and sequence analysis were performed at CUGI  
 (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).  
 The sequence has been trimmed to remove vector sequence  
 and contains a minimum of 100 bases of phred value 20 or  
 above. For more details on library preparation and  
 sequence analysis see  
 http://www.genome.clemson.edu/projects/barley. To order  
 this clone see http://www.genome.clemson.edu/orders  
 Also see Close TJ, Wing R, Kleinofs A, Wise R (2001)  
 Genetically and physically anchored EST resources for  
 barley genomics. Barley Genetics Newsletter 31:29-30.  
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

BASE COUNT 220 a 146 c 198 g 236 t 1 others  
 ORIGIN

Query Match 13.0%; Score 130; DB 9; Length 801;  
 Best Local Similarity 96.0%; Pred. No. 7.6e-20;  
 Matches 144; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 43 aggatattgatattcattcagctcgtgatatg-gctcaactcttcgattgacgtgaca 101  
 Db 291 AGGATATGATGATTTTCATGGCTGCTGATAGACCTTCAACCCCTGCATTGATCGTGCA 350

QY 102 tagcattacataaataagaccagctgtcaccatgggttagtggtagaactatctta 161  
 Db 351 TAGCATTACATAAATAAGACGAGCTGTACCATGGGTTAGGTGCGAAGCTATCTTA 410

QY 162 acttcattggaatgagttgggcatctg 191  
 Db 411 ATTTCATGGGAATGAGTTGGGCATCTTG 440

RESULT 9

LOCUS BF065047 825 bp mRNA linear EST 23-OCT-2001  
 DEFINITION HV\_CEB0022L05f Hordeum vulgare seedling green leaf EST library  
 HVCNMA0005 (Blumeria challenged) Hordeum vulgare cDNA clone  
 HV\_CEB0022L05f, mRNA sequence.  
 BF065047  
 VERSION BF065047.1 GI:10841686  
 KEYWORDS EST.  
 SOURCE barley.  
 ORGANISM Hordeum vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
 ; Triticeae; Hordeum.  
 1 (bases 1 to 825)  
 Wing,R., Close,T.J., Kleinofs,A., Wise,R., Wei,F., Begum,D.,  
 Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi  
 ,D.W., Fenton,R.D., Oates,R. and Main,D.  
 Development of a genetically and physically anchored EST resource  
 for barley genomics: Blumeria infected incompatible (Mla6) seedling  
 leaf cDNA library  
 Unpublished (2001)  
 CONTACT: Wing RA  
 CLEMSON UNIVERSITY GENOMICS INSTITUTE  
 CLEMSON UNIVERSITY  
 100 JORDAN HALL, CLEMSON, SC 29634, USA  
 TEL: 864 656 7288  
 FAX: 864 656 4293  
 EMAIL: rwing@clermson.edu  
 Total hg bases = 594  
 Seq primer: AATTAACTCTCACTAAAGCG  
 High quality sequence stop: 732.  
 Location/Qualifiers  
 1..825  
 /organism="Hordeum vulgare"  
 /cultivar="C16151 (Mla6)"  
 /db\_xref="taxon:4513"  
 /clone\_lib="HV\_CEB0022L05f"  
 HVCNMA0005 (Blumeria challenged)"  
 /clone\_lib="Hordeum vulgare seedling green leaf EST  
 library HVCNMA0005 (Blumeria challenged)"  
 /tissue\_type="seedling green leaf"  
 /lab\_host="SOLR"  
 /note="Vector: lambdaZAP; Site\_1: EcoRI; Site\_2: XhoI;  
 C.I. 16151 (Mla6) plants were greenhouse grown in the R  
 Wise lab at Iowa State University, Ames, IA; 7 day old  
 green seedlings were challenged with isolate 5874 (AvrMla6  
 ) of Blumeria graminis f. sp. hordei, and leaves were  
 harvested 20 and 24 hr post-inoculation and snap frozen;  
 uninoculated leaves were harvested 20 hr post-inoculation  
 (Wei, Wise). In the TJ Close lab at the University of  
 California, Riverside, total RNA was prepared from each  
 sample pool, equal quantities of all three RNA pools were



Qy 185 catcccg 191  
|||||  
Db 234 CATCTG 240

RESULT 12  
BE498533  
LOCUS  
DEFINITION WHE0971.G03\_M05ZS wheat pre-anthesis spike cDNA library Triticum  
ACCESSION BE498533  
VERSION BE498533.1 GI:9697150  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Triticum.  
1 (bases 1 to 494)  
REFERENCE  
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han  
P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,  
Seaton,C.L. and Tong,J.C.  
The structure and function of the expressed portion of the wheat  
genomes - Pre-anthesis spike cDNA library  
Unpublished (2000)  
JOURNAL  
COMMENT US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105395773  
Fax: 5105395818  
Email: oanderson@w.usda.gov  
Sequence have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: Stratagene SK primer.  
Location/Qualifiers  
1..494  
/organism="Triticum aestivum"  
/cultivar="Chinese Spring"  
/db\_xref="taxon:4565"  
/clone="WHE0971.G03\_M05"  
/clone\_lib="Wheat pre-anthesis spike cDNA library"  
/tissue\_type="Spike before anthesis"  
/dev\_stage="Adult plant"  
/lab\_host="E. coli SOLR"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
Site\_1: EcoRI; Site\_2: XhoI; Plants were grown in the  
greenhouse. Whole spike with awns trimmed, white, green  
and yellow anther were collected and total RNA, and  
poly(A) RNA were prepared, a cDNA library was made, and  
the cDNA clones were in vivo excised to give pluescript  
phagemids in the IV Close lab (Choi, Close, Fenton) at  
the University of California, Riverside. Plasmid DNA  
preparations and DNA sequencing were performed in the OD  
Anderson lab (all other authors)."

BASE COUNT 129 a 106 c 135 g 124 t

ORIGIN

Query Match 11.4%; Score 114.2; DB 10; Length 494;  
Best Local Similarity 97.5%; Pred. No. 3, 2e-16;  
Matches 116; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 883 gttatgacatctgagcaccagatgttcaacggaacatgaggaataagtgatcat 942  
|||||  
Db 85 gttatgacatctgagcaccagatgttcaacggaacatgaggaataagtgatcat 144

Qy 943 ccccaaaaggagattgttattgtttcaactccactgagcaatgctttttg 1001  
|||||  
Db 145 cttcgaaaggagattgttattgtttcaactccactgagcaatgctttttg 203

RESULT 13

AV944141/c  
LOCUS  
DEFINITION AV944141 688 bp mRNA linear EST 18-JAN-2002  
AV944141 K. Sato unpublished cDNA library, strain H602 adult,  
heading stage top three leaves Hordeum vulgare subsp. spontaneum  
cDNA clone ban20e24 3', mRNA sequence.  
AV944141  
ACCESSION AV944141.1 GI:18239938  
VERSION  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. spontaneum.  
ORGANISM Hordeum vulgare subsp. spontaneum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
1 (bases 1 to 688)  
REFERENCE  
AUTHORS Sato,K., Saitoh,D. and Takeda,K.  
TITLE Barley EST sequencing project in NIG and Okayama Univ  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-1  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp.  
Location/Qualifiers  
1..688  
/organism="Hordeum vulgare subsp. spontaneum"  
/strain="H602"  
/db\_xref="taxon:77009"  
/clone="bah20e24"  
/clone\_lib="K. Sato unpublished cDNA library, strain H602  
adult, heading stage top three leaves"  
/tissue\_type="top three leaves"  
/dev\_stage="adult, heading stage"

BASE COUNT 184 a 180 c 145 g 177 t 2 others

ORIGIN

Query Match 11.4%; Score 114.2; DB 9; Length 688;  
Best Local Similarity 97.5%; Pred. No. 3, 4e-16;  
Matches 116; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 883 gttatgacatctgagcaccagatgttcaacggaacatgaggaataagtgatcat 942  
|||||  
Db 612 gttatgacatctgagcaccagatgttcaacggaacatgaggaataagtgatcat 553

Qy 943 ccccaaaaggagattgttattgtttcaactccactgagcaatgctttttg 1001  
|||||  
Db 552 cttcgaaaggagattgttattgtttcaactccactgagcaatgctttttg 494

RESULT 14

AV921312 692 bp mRNA linear EST 18-JAN-2002  
AV921312 K. Sato unpublished cDNA library, cv. Haruna Nijo  
germination shoots Hordeum vulgare subsp. vulgare cDNA clone  
bags23a09 3', mRNA sequence.  
AV921312  
ACCESSION AV921312.1 GI:18217091  
VERSION  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare.  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
1 (bases 1 to 692)  
REFERENCE  
AUTHORS Sato,K., Saitoh,D. and Takeda,K.  
TITLE Barley EST sequencing project in NIG and Okayama Univ  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadasu Shin-1  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.  
 Location/Qualifiers

# FEATURES

Source

1. 692

/organism="Hordeum vulgare subsp. vulgare"  
 /cultivar="Haruna Nijo"  
 /db\_xref="taxon:112509"  
 /clone="bags23a09"  
 /clone\_id="K. Sato unpublished cDNA library, cv. Haruna  
 Nijo germination shoots"  
 /tissue\_type="shoots"  
 /dev\_stage="germination"  
 BASE COUNT 183 a 184 c 147 g 176 t 2 others  
 ORIGIN

Query Match 11.4%; Score 114.2; DB 9; Length 692;  
 Best Local Similarity 97.5%; Pred. No. 3.4e-16;

Matches 116; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 883 gttttagacatctggaccagatgtttcagcgaacatagagaataagtgatcat 942

DB 580 GTTTATGACATCTGAGCACCAGTATGTTCTCGAAACATGAGAGATAGAGTGATCAT 521

OY 943 cctcaaaagagagagattgttattgttttcaacttcacactgagacaatagcttttg 1001

DB 520 CTTGCAAGAGAGATTGGTATTGTTTTCACCTCCAGCAGCAATAGCTTTTGG 462

## RESULT 15

LOCUS B1948391 728 bp mRNA linear EST 19-OCT-2001

DEFINITION HVSMEL0009F19f Hordeum vulgare spike EST library HVCDNA0012  
 (Fusarium infected) Hordeum vulgare cDNA clone HVSMEL0009F19f, mRNA  
 sequence.

ACCESSION B1948391 GI:16288749

VERSION B1948391

KEYWORDS EST.

SOURCE barley.

ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 ; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 728)  
 Wing, R., Muehlbauer, G.J., Close, T.J., Kleinhofs, A., Wise, R., Heinlen,  
 S., Begum, D., Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T.,  
 Simmons, J., Fenton, R.D., Malatrasi, M., Choi, D.W., Oates, R. and Main,  
 D.

TITLE Development of a genetically and physically anchored EST resource  
 for barley genomics: Fusarium infected Morex spike cDNA library

JOURNAL Unpublished (2001)

COMMENT Contact: Wing RA  
 Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 639

Seq primer: AATTAACTCTCCTAAAGG

High quality sequence stop: 680.

Location/Qualifiers

1. 728

/organism="Hordeum vulgare"  
 /cultivar="Morex"  
 /db\_xref="taxon:4513"  
 /clone="HVSMEL0009F19f"  
 /clone\_id="Hordeum vulgare spike EST library HVCDNA0012  
 (Fusarium infected)"  
 /tissue\_type="Spike"  
 /lab\_host="TJC121"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:

XhoI; Plants were grown at the University of Minnesota in the Gf Muehlbauer lab; spikes were harvested and snap frozen at 0, 1, 2, 3, 4, 5, 6, and 8 days after Fusarium graminearum inoculation (Heinlen). In the Tj Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all eight RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Choi, Fenton, Malatrasi ). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/99pages/bgn/31/cover.html)"

BASE COUNT 192 a 151 c 188 g 192 t 5 others  
 ORIGIN

Query Match 11.4%; Score 114.2; DB 10; Length 728;  
 Best Local Similarity 97.5%; Pred. No. 3.4e-16;

Matches 116; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 883 gttttagacatctggaccagatgtttcagcgaacatagagaataagtgatcat 942

DB 62 GTTTATGACATCTGAGCACCAGTATGTTCTCGAAACATGAGAGATAGAGTGATCAT 121

OY 943 cctcaaaagagagagattgttattgttttcaacttcacactgagacaatagcttttg 1001

DB 122 CTTGCAAGAGAGATTGGTATTGTTTTCACCTCCAGCAGCAATAGCTTTTGG 180

Search completed: September 12, 2002, 08:10:33  
 Job time: 16663 sec



---

